
Desarrollo de un chatbot de asesoría en casos de acoso escolar



Trabajo de Fin de Grado
Curso 2018–2019

Autor

María Soledad Espinosa Ruiz

Director

Pablo Gervás Gómez-Navarro

Grado en Ingeniería Informática

Facultad de Informática

Universidad Complutense de Madrid

Desarrollo de un chatbot de asesoría en casos de acoso escolar

Trabajo de Fin de Grado en Ingeniería Informática
Departamento de Ingeniería del Software e Inteligencia
Artificial

Autor

María Soledad Espinosa Ruiz

Director

Pablo Gervás Gómez-Navarro

Convocatoria: *Septiembre 2019*

Calificación:

Grado en Ingeniería Informática
Facultad de Informática
Universidad Complutense de Madrid

19 de septiembre de 2019

Dedicatoria

A mi madre, por enseñarme a luchar hasta el final y a amar sin medida.

A Inés, que solo con abrir los ojos, todo lo ocupa.

A ti, la casualidad de mi vida, la más grande.

Agradecimientos

En primer lugar me gustaría agradecer a mi director, Pablo Gervás, por aceptar y enriquecer mi propuesta desde el primer día, por asesorarme y guiarme cuando me perdía en el camino y, sobre todo, por sus palabras de ánimo y su comprensión en las innumerables ocasiones en las que las he necesitado.

También me gustaría agradecer a otros muchos profesores que me han motivado a seguir adelante en esta complicada aventura:

a Carlos León, por el esfuerzo y, sobre todo, la infinita sabiduría,

a Gonzalo Méndez, por el sentido del humor y los buenos consejos,

a Jesús Correas, por escucharme y comprenderme,

y a todos esos profesores que hacen que su trabajo cambie la vida de las personas que se cruzan con ellos.

Un agradecimiento especial merecen también Andrea Fuentes y Rocío Arenillas, las psicólogas que se han atrevido a hablar con Otto y a buscarle las cosquillas. No solo son grandes profesionales, sino que además, son mejores personas, de las que tampoco quedan muchas.

También me gustaría dar las gracias al departamento de Ingeniería del Software e Inteligencia Artificial, por los valiosos recursos que me han prestado para el desarrollo técnico de OttoBot y a todos los profesionales que hacen posible el funcionamiento de la facultad.

Por último, gracias a mi familia, que me apoya y me soporta más de lo que seguramente merezco, a mis compañeras de carrera, que son tres tesoros de incalculable valor y a mi suegra, por no perder la confianza en mí.

Abstract

Bullying is an issue that many children and teenagers face at some point in their lives, either by being victimized, by witnessing it or even by bullying others. Research has shown that being victimised at early stages of life can have a tremendous negative impact on mental and physical development. This project aims to develop a system that can be used as a tool for future psychologists in their training for school bullying therapy. Using the features offered by AIML, OttoBot was developed as a prototype model of in-therapy conversation by a 13 year-old boy.

Based on a review of the literature, the emotional model of OttoBot was created as a set of variables that could increase, decrease, and combine to map to the different behaviours of OttoBot. In order to test the features of the system, a formative evaluation was carried out by experts of the field. The four phase assessment combined conversations with OttoBot as well as formative evaluation questionnaires in which the experts had not only to rate but also to comment on the key features of OttoBot. The results indicate that, with the appropriate optimizations, OttoBot might be a powerful tool for the training of future psychologists. Further research and work is needed to improve and enrich the possibilities of OttoBot as a training tool.

Keywords

chatbot, bullying, AIML, conversational agent, dialogue system, therapy, chat.

Resumen

El acoso escolar es un problema al que muchos niños y adolescentes se enfrentan en algún momento de sus vidas, ya sea por ser víctimas, por presenciarlo o, incluso, por intimidar a otros. Investigaciones recientes demuestran que ser víctima de acoso escolar durante las primeras etapas de la vida puede tener un tremendo impacto negativo en el desarrollo mental y físico de niños y jóvenes. Este proyecto tiene como objetivo desarrollar un sistema que pueda usarse como una herramienta para futuros psicólogos en su formación práctica para terapias de acoso escolar. Utilizando las funcionalidades ofrecidas por AIML, OttoBot fue desarrollado como un prototipo de modelo de conversación de un niño de 13 años en terapia.

Tras una revisión de la literatura, el modelo emocional de OttoBot se creó como un conjunto de variables que podrían aumentar, disminuir y combinarse para establecer los diferentes comportamientos de OttoBot. Con el fin de probar las características del sistema, expertos en la materia realizaron una evaluación formativa. En las cuatro fases de la evaluación se combinaron conversaciones con OttoBot así como cuestionarios de evaluación formativa en los que los expertos no solo tuvieron que calificar sino también comentar las características clave de OttoBot. Los resultados indican que, con las optimizaciones apropiadas, OttoBot podría ser una herramienta de gran utilidad para la formación práctica de futuros psicólogos. Se necesita más investigación y trabajo para mejorar y enriquecer las posibilidades de OttoBot como herramienta de capacitación.

Palabras clave

chatbot, bullying, AIML, agente conversacional, sistema de diálogo, terapia, acoso escolar.

Contents

1. Introducción	1
1.1. Motivación	2
1.2. Objetivos	3
1.3. Plan de trabajo	3
1.4. Estructura del proyecto	5
1. Introduction	7
1.1. Motivation	8
1.2. Objectives	9
1.3. Work Plan	9
1.4. Project Structure	10
2. State of the Art	13
2.1. Bullying	13
2.1.1. The concept of bullying	13
2.1.2. Evolution of bullying	15
2.2. Chatbot technology as a tool for bullying	17
2.2.1. History of chatbots	17
2.2.2. Different chatbot configurations	22
2.2.3. Previous work on bullying using chatbot technology	24
2.3. Other technological approaches to bullying	26
3. Architecture and Implementation	27

3.1. Overview	27
3.2. User interaction modules	27
3.2.1. User Interaction	28
3.2.2. Google Hangouts Chat API	30
3.3. Internal dialog-management modules	33
3.3.1. User Message and Language Processing	33
3.3.2. Context, Retrieval-Based Model, and Responses	34
3.3.3. Data Storage and Analytics	37
3.3.4. Bot engine as a UNIX service with <code>systemd</code>	37
4. A model of the bot's behaviour	41
4.1. Adaptation of user input	42
4.2. A computational model of a bullied subject's state	42
4.2.1. Variables representing subject-therapist relationship	42
4.2.2. Implementation of the subject responses according to the current emotional state	44
4.3. Question set acquisition	47
4.3.1. Methodology	47
5. Formative evaluation	51
5.1. Evaluation process	52
5.1.1. Formative evaluator profiles	53
5.2. Analysis of results	53
6. Discussion	61
7. Conclusions and Future Work	63
7.1. Future work	63
7. Conclusiones y trabajo futuro	65
7.1. Trabajo Futuro	65
A. Survey: <i>Reescritura de preguntas</i>	67
A.1. Distribution Message	68
B. Factores del CUVE3-ESO e ítems que los componen	69

C. Evaluation conversations: Subject 1	73
C.1. Phase 1: First conversation	73
C.2. Phase 3: Second conversation	75
D. Evaluation conversations: Subject 2	77
D.1. Phase 1: First conversation	77
D.2. Phase 3: Second conversation	79
Bibliography	81

List of figures

2.1. Venn's diagram representig the model proposed by Olweus (1993).	14
2.2. Example of conversation with ELIZA.	18
2.3. Example of conversation with PARRY.	19
2.4. Example of conversation with Jabberwacky.	20
2.5. Example of conversation with Dr.Sbaitso.	21
2.6. Example of conversation with Mitsuku.	22
2.7. Example of conversation with Siri.	23
3.1. Modules that conform the system's main architecture.	28
3.2. Number of daily active users in the different instant messaging apps in Google Play in Spain (April 2018).	29
3.3. Market share for operating systems in Spain.	30
3.4. Example of simple conversation with OttoBot.	31
4.1. Ranges of possible values for the context variables.	44
5.1. Responses to question 1 given by the study subjects.	54
5.2. Responses to question 2 given by the study subjects.	55
5.3. Responses to question 3 given by the study subjects.	56
5.4. Responses to question 4 given by the study subjects.	57
5.5. Responses to question 5 given by the study subjects.	58
5.6. Responses to question 6 given by the study subjects.	59
5.7. Assessment questionnaire.	60

List of tables

1.1. Fechas establecidas para cada una de las fases del plan de trabajo	4
1.1. Established deadlines for each of the work plan phases.	10
3.1. AIML tags description table.	35
4.1. User's input influence on bot's trust, acceptance, and commitment	44
4.2. CUVE ³ -ESO factors and items that compose it (Álvarez et al. (2013)).	48

Capítulo 1

Introducción

Aunque el *bullying* siempre ha existido, en los últimos años se ha convertido en un problema muy serio. Muchos jóvenes se enfrentan al acoso escolar en algún momento de su vida, ya sea como víctimas, como testigos o incluso como perpetradores.

“Una persona está sufriendo bullying cuando está expuesta de manera repetida a acciones negativas por parte de uno o más de los otros estudiantes” (Olweus (1993)). Monks and Smith (2006) define la palabra *bullying* como una relación caracterizada por una agresión continuada y con una asimetría de poder: un subtipo de acoso que puede parecer injusto a terceros y que puede tener efectos serios sobre la víctima.

Independientemente de las cuestiones definitorias, existe unanimidad en la creencia de que la prevención del acoso escolar y la reducción de la victimización podría tener un gran impacto en las consecuencias para las víctimas. Hoy en día, este tipo de comportamientos destructivos repetitivos pueden ocurrir en el aula, en el campo de fútbol o incluso en las redes sociales, y los niños y los jóvenes parecen ser especialmente propensos a sentir el dolor causado por ser víctimas. De hecho, numerosas investigaciones sugieren que los estudiantes acosados sufren una disminución de la autoestima, faltan a clase a menudo (para evitar la victimización) y se sienten inseguros en el entorno escolar. Los resultados han demostrado que las víctimas de acoso pueden experimentar sentimientos de rechazo, soledad y, en casos extremos, tienen un mayor riesgo de suicidio (Esbensen and Carson (2009)).

En contra de la creencia tradicional popular de que el *bullying* es algo típico del comportamiento juvenil, los estudios llevados a cabo en las últimas décadas han demostrado que ser víctima de acoso escolar en la niñez o adolescencia puede tener un tremendo impacto en el desarrollo mental, físico o académico.

En los últimos años se han desarrollado muchas políticas anti-bullying y programas de prevención con la finalidad de mitigar los efectos negativos que el acoso escolar puede tener en niños y adolescentes. Aunque estos han estado generalmente enfocados en los efectos sobre las víctimas, también se ha prestado atención a las posibles consecuencias sociales que el bullying puede tener sobre aquellos que lo cometen e incluso sobre aquellos que participan como víctima y perpetrador. En sus investigaciones sobre las consecuencias del bullying, Rigby (2003) las categoriza en cuatro tipos: (1) malestar psicológico, (2) mala adaptación social, (3) trastornos psicológicos y (4) malestar físico.

Esta situación ha provocado el desarrollo de diversos programas de prevención en escuelas en todo el mundo (por ejemplo, Farrington (1993) y Olweus et al. (1999)), que intentan reestructurar el ambiente escolar existente para reducir las oportunidades de que ocurra el acoso así como para mejorar las habilidades sociales de las víctimas, la empatía de los acosadores y la supervisión por parte de los adultos en los centros escolares.

1.1. Motivación

Además de las técnicas de prevención y herramientas mencionadas, también existe la necesidad de crear nuevos programas que hagan referencia al tratamiento y las opciones de terapia. Como ya establece Rigby (2003), *“se pueden considerar muchas estrategias o tratamientos para reducir las oportunidades de que un niño se vea involucrado en casos de acoso escolar. Éstas incluyen ayudar a los niños victimizados a desarrollar habilidades de asertividad autoprotectora y trabajar de manera terapéutica con los niños que cometen bullying para establecer una mayor conciencia de las consecuencias de su comportamiento antisocial.”*

Teniendo en cuenta los antecedentes descritos anteriormente, se ha reconocido una creciente necesidad de nuevos profesionales bien capacitados que puedan contribuir a disminuir las consecuencias del acoso escolar en niños y jóvenes. Esta necesidad podría mitigarse con la existencia de un software automatizado que ayude a los psicólogos a mejorar sus habilidades terapéuticas.

El gran desarrollo tecnológico experimentado en las últimas décadas ha contribuido a un cambio cultural en lo que respecta a las relaciones y a las comunicaciones sociales. Se han desarrollado varias soluciones tecnológicas utilizando los beneficios que ofrecen los entornos virtuales para abordar el problema del acoso escolar. El estudio realizado por Nocentini et al. (2015) ofrece una revisión sistemática de la intervención contra el acoso escolar mediante el uso de tecnologías de la información, analizando las características de los principales programas y la evidencia de su eficacia y concluyendo que

las tecnologías digitales son una característica de la sociedad actual y su uso destaca especialmente durante la infancia y la adolescencia. Estas razones motivan al presente proyecto a intentar crear una de esas herramientas.

1.2. Objetivos

El objetivo de este proyecto es crear una herramienta que ayude a futuros psicólogos en el entrenamiento para la terapia con niños y jóvenes en general. En particular, se han establecido los siguientes objetivos:

- Identificación de las características de la actitud del paciente que más influyen en el resultado de la terapia.
- Reconocimiento de las preguntas más comunes a la hora de determinar si un niño o adolescente sufre acoso escolar.
- Creación de un modelo computacional de diálogo basado en las características mencionadas en los puntos anteriores.
- Implementación de dicho modelo en una aplicación basada en chats.
- Despliegue de la aplicación como un servicio en línea.
- Integración de dicho servicio con la API de Google Hangouts Chat.

1.3. Plan de trabajo

Con el fin de cumplir con los objetivos establecidos en la Sección 1.2, éstos se han dividido en tareas más pequeñas y detalladas. La siguiente lista describe las principales funcionalidades que se van a desarrollar. Además, en la Tabla 1.1 se encuentran las fechas en las que dichas funcionalidades han de ser implementadas.

- **Fase 1:** Repaso exhaustivo de la literatura relacionada con la historia del acoso escolar, los agentes conversacionales y las herramientas tecnológicas desarrolladas con la finalidad de abordar el problema del acoso escolar para construir una sección que englobe el estado actual de la cuestión.
- **Fase 2:** Diseño del modelo computacional. Proceso de toma de decisiones en el que se establezcan las principales características funcionales del bot, como por ejemplo los tipos de respuesta y las actitudes que caracterizarán a OttoBot.

Fase	Descripción	Fecha
1	Investigación	1 - 30 Nov. 18
2	Diseño del modelo	1 - 15 Dic. 18
3	Diseño de la arquitectura	16 Dic. 18 - 15 Ene. 19
4	Desarrollo de <i>software</i> : Bot	16 Ene. - 28 Feb. 19
5	Desarrollo de <i>software</i> : Infraestructura	1 Mar. - 10 Mar. 19
6	Resolución de problemas	11 Mar. - 31 Mar. 19
7	<i>Testing</i>	1 Apr. - 20 Abr. 19
8	Documentación: capítulos 3 y 4	21 Abr. - 15 May. 19
9	Evaluación cualitativa	15 May. - 30 Jun. 19
10	Documentación: capítulos 5 y 6	1 Jul. - 31 Jul. 19
11	Documentación: capítulo 7	1 Ago. - 31 Ago. 19

Tabla 1.1: Fechas establecidas para cada una de las fases del plan de trabajo

- **Fase 3:** Diseño de la arquitectura y elección de las principales tecnologías que se van a utilizar como los lenguajes de programación, la tecnología de chatbot, la infraestructura de chat, el entorno de desarrollo, etc.
- **Fase 4:** Desarrollo de software. Desarrollo del bot en JavaScript e integración con el procesador AIML escogido: AIMLInterpreter.
- **Fase 5:** Desarrollo de software. Integración de la infraestructura de la aplicación del chatbot en JavaScript y AIML con la API de Google Hangouts Chat. Despliegue de OttoBot en un Servidor Ubuntu y creación de los servicios como `systemd`.
- **Fase 6:** Resolución de problemas del modelo y cuestiones correspondientes a la arquitectura, como la necesidad de un estado y un contexto del bot.
- **Fase 7:** Análisis y mejoras de la aplicación (*testing*).
- **Fase 8:** Desarrollo de la documentación correspondiente a la arquitectura y al modelo de OttoBot, concretamente: elaboración de los capítulos 3 y 4.
- **Fase 9:** Evaluación cualitativa hecha por un experto en psicología con experiencia en terapia infantil.
- **Fase 10:** Resultados, análisis y documentación de las evaluaciones cualitativas, concretamente: capítulos 5 y 6.
- **Fase 11:** Reconocimiento de las limitaciones y finalización de las conclusiones y la documentación.

1.4. Estructura del proyecto

En esta sección se presenta al lector la estructura que sigue el presente documento. Para comenzar, la literatura más relevante sobre los temas de bullying, agresión y tecnología de chatbot ha sido investigada y documentada en el Capítulo 2. Seguido de una descripción técnica de los módulos que forman la arquitectura de la aplicación y una descripción funcional del modelo de conversación desarrollado, como se puede encontrar en los Capítulos 3 y 4, respectivamente. Posteriormente, en el Capítulo 5, se presentan los principales resultados de la evaluación del trabajo realizado, seguido de una discusión crítica en el Capítulo 6. Finalmente, el Capítulo 7 resume las conclusiones así como las nuevas áreas de investigación provocadas por los desarrollos de este proyecto.

Chapter 1

Introduction

Even though it has always existed, bullying has become a very serious issue in the later years. Many children and teenagers face bullying at some point in their lives, either by being victimized, by witnessing it or even by bullying others.

“A person is being bullied or victimised when he or she is exposed, repeatedly over time, to negative actions on the part of one or more other students” Olweus (1993). The word bullying as Monks and Smith (2006) define it is a relationship characterised by continued aggression and with a power asymmetry – a picking on or harassment, which can appear unfair to onlookers and which can have serious effects for those who are victims.

Regardless of definitional concerns, there is unanimity in the belief that preventing bullying and reducing victimization might have a great impact on the consequences for victims. Nowadays, this kinds of repetitive destructive behaviours can occur in the classroom, on the soccer field, or on a social networking website and, children and young people seem to be especially prone to feel the pain caused by being victimized. In fact, there is substantial research suggesting that bullied students suffer from a lowering of self-esteem, to being frequently absent from school (to avoid victimization), and to feeling unsafe and insecure in the school setting. Research has demonstrated that victims of bullying can experience feelings of rejection, loneliness, and in extreme cases, are at increased risk for suicide(Esbensen and Carson (2009)).

Contrary to the traditional belief that bullying is a typical behaviour during childhood and young age, research conducted in the las decades demonstrated that experiencing bullying victimization at a young age can have a tremendous impact on mental, physical or academical development.

In recent years many anti-bullying policies and prevention programs have been developed in order mitigate the negative effects that bullying might

have on children and young people. Even though they have been mainly focused on the effects on the victims of bullying, some attention has also been given to the possible social consequences for those who bully others and also to those involved in it as both victim and perpetrator. In its research on the consequences of bullying, Rigby (2003) categorizes them into four types: (1) low psychological wellbeing, (2) poor social adjustment, (3) psychological distress, and (4) physical unwellness.

This situation has led to the development of several prevention programs in schools around the globe (e.g. Farrington (1993) and Olweus et al. (1999)), which attempt to restructure the existing school environment to reduce opportunities and rewards for bullying behaviour as well as to improve the social skills of the victim, the empathy of the bully, and the adult supervision in school environments.

1.1. Motivation

In addition to prevention techniques and tools, there is also a need for new programs regarding treatment and therapy options. As stated by Rigby (2003) in its study, *“various strategies or treatments may be considered to reduce the chances of a child’s further involvement in bullying that may worsen the condition. These include assisting victimised children to develop self-protective assertiveness skills and working therapeutically with bullying children to establish a greater awareness of the consequences of their antisocial behaviour.”*

Given the previously described background, an increasing need for well trained professionals that can contribute to lessen the consequences that experimenting bullying can have in children and young people has been acknowledged. This need can be mitigated with the existence of automated software that can help psychologists to improve their therapy skills.

The great development that communication technologies have had in the past decades has contributed to a cultural change in what regards to social and relational communication domains. Several technological solutions have been developed using the benefits offered by virtual environments in order to address the issue of bullying. A study carried out by Nocentini et al. (2015) presents a systematic review on anti-bullying ICT-mediated intervention, analyzing the characteristics of the main programs and the evidence of their effectiveness and concluding that digital technologies are a feature of nowadays society and their use is particularly salient during childhood and adolescence. These reasons motivate the present project to attempt to create one of such tools.

1.2. Objectives

The aim of the present project is to create a tool that will help psychologists in their training for therapy of bullied children and youth in general. In particular, the following list of objectives has been established:

- Identification of the most relevant features of the patient's attitude that can influence the outcome of therapy
- Recognition of the most common questions in order to determinate if a child or teenager is being bullied
- Creation of a computational dialogue model based on the previously stated features
- Implementation of the model in a chat-based application
- Deployment of the application as an online service
- Integration of the service with Google Hangouts Chat API

1.3. Work Plan

In order to accomplish the objectives described in Section 1.2, they need to be divided into smaller, more detailed tasks. The following list describes the main features to be developed and the dates by which those features must be implemented are detailed in Table 1.1.

- **Phase 1:** Exhaustive revision of relevant literature related to the history of bullying, conversational agents, and technological tools developed to address the issue of bullying in order to build the state-of-the-art section of the topic.
- **Phase 2:** Computational dialog model design. Decision-taking process in order to establish the main functional features of the bot, such as the kinds of responses and attitudes which would characterise OttoBot.
- **Phase 3:** Architecture design and election of main technologies to use such as programming language, bot technology, chat infrastructure, developing environment, etc.
- **Phase 4:** Software development. Bot development in JavaScript and integration with the chosen AIML parser, AIMLInterpreter.

Phase	Description	Date
1	Research	1 - 30 Nov. 18
2	Model Design	1 - 15 Dec. 18
3	Architecture Design	16 Dec. 18 - 15 Jan. 19
4	Software Development: Bot	16 Jan. - 28 Feb. 19
5	Soft. Development: Infrastructure	1 Mar. -10 Mar. 19
6	Troubleshooting	11 Mar. - 31 Mar. 19
7	Testing	1 Apr. - 20 Apr. 19
8	Documentation: Chapters 3 & 4	21 Apr. - 15 May 19
9	Qualitative Evaluation	15 May - 30 Jun. 19
10	Documentation: Chapters 5 & 6	1 Jul. - 31 Jul. 19
11	Documentation: Chapter 7	1 Aug. - 31 Aug. 19

Table 1.1: Established deadlines for each of the work plan phases.

- **Phase 5:** Software development. Infrastructure integration of the bot application in JavaScript and AIML with the Google Hangouts Chat API. Deployment of OttoBot in Ubuntu Server and systemd services creation.
- **Phase 6:** Troubleshooting. Problem solving of model and architecture issues such as the need for a state and a context of the bot.
- **Phase 7:** Testing and improvement of the application.
- **Phase 8:** Development of the documentation corresponding to the architecture and the model of OttoBot, namely development of Chapters 3 and 4.
- **Phase 9:** Qualitative evaluation made by an expert psychologist with children therapy experience.
- **Phase 10:** Evaluation result, analysis and documentation, namely completion of Chapters 5 and 6.
- **Phase 11:** Limitation acknowledgement, conclusions and documentation completion.

1.4. Project Structure

This section presents the reader with the structure to be found in the present document. Initially, the most relevant literature on the topics of

bullying, aggression, and chatbot technology has been researched and documented in Chapter 2. Followed up by a technical description of the modules that form the application's architecture and a functional description of the conversational model developed, as can be found in Chapters 3 and 4, respectively. Afterwards, in Chapter 5, the main results of the evaluation of the work done are presented followed by with a critical discussion in Chapter 6. Ultimately, Chapter 7 summarizes the conclusions as well as the new areas of investigation prompted by developments of this project.

Chapter 2

State of the Art

In order to understand not only the purpose but also the motivation of the present project, it is important to explain, on the one hand, the meaning and evolution of the concept of bullying in our society and, on the other hand, the history and different techniques of speech conversation systems or chatbots, exploring the different configurations that have been and can be done using this technology. The final part of this section will include a review of all the related work regarding chatbots as a tool for treating bullying as well as other technological approaches to the issue along the history of artificial intelligence.

2.1. Bullying

2.1.1. The concept of bullying

Bullying is clearly related to the concepts of violence and aggression. In order to get an accurate definition of bullying, it is necessary to explore first these two concepts. As Baron and Richardson (2004) suggest, a definition of aggression would be one that included the involvement of intention as well as the actual delivery of harm to others. Carlson et al. (1989) states that human aggression represents the intentional attempt to harm others. Although the term needs further revision, it can be agreed that aggression and aggressive behaviour include intention and infliction of harm.

With regard to violence, the OED online (1989c) defines it as the deliberate exercise of physical force against a person or property. Olweus (1993) states that violence is clearly a subcategory of aggression, and defines it as an act where the perpetrator inflicts injury upon the victim. The previous definition presents itself as very interesting to the matter of this project, be-

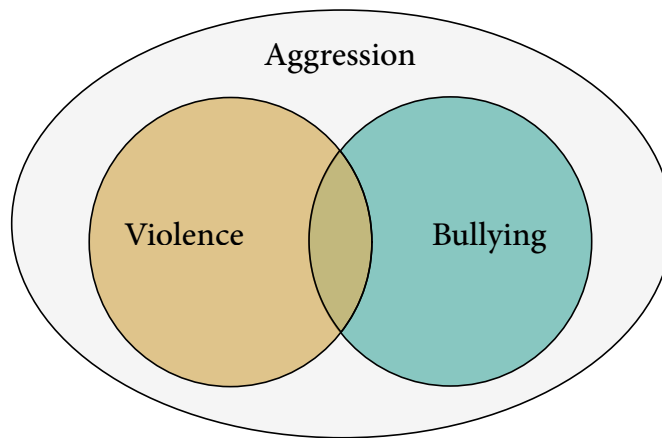


Figure 2.1: Venn's diagram representig the model proposed by Olweus (1993).

cause it also relates bullying, and aggression with the term. Olweus (1993) states that violence and bullying are subcategories of aggression or aggressive behaviour that might as well overlap. The described relations are depicted in Figure 2.1.

Etymologically the term bullying comes from the verb to bully, whose usage traces back to the early 1700s. OED online (1989a), meaning “*to treat in an overbearing manner; to intimidate*”. OED online (1989b) defines bullying as the “*action of the verb to bully: overbearing insolence, personal intimidation, petty tyranny. Often used with reference to schoolboy life*”. One of the most acknowledged definitions of bullying is the one given by Olweus (1993), which states that “*A person is being bullied or victimised when he or she is exposed, repeatedly over time, to negative actions on the part of one or more other students*”. Monks and Smith (2006) refer also to the effects that bullying causes on the victim as part of the definition of bullying itself: “*bullying is thus a relationship characterised by continued aggression and with a power asymmetry – a picking on or harassment, which can appear unfair to onlookers and which can have serious effects for those who are victims*”.

Despite the fact that there is no universal definition of bullying, there is some consensus in that bullying always includes:

- the figures of the abuser (or bully) and the victim
- an intent to harm either physically or mentally
- repetition over time
- imbalance in power

2.1.2. Evolution of bullying

The study of the evolution of bullying presents the researchers with three different dimensions that need to be explored in order to fully understand how bullying has changed over time.

First of all, there are cultural facts that need to be considered, as the traditional and modern understanding of aggressive behaviour has been changing differently across the globe. As stated by Carey (2003), bullying has been researched since the early 1960s, and was first researched on a large scale in the early 1970s by Dr. Olweus in Sweden. A decade later, the first national campaign promoted by a Ministry of Education was initiated in Norway, resulting in the aforementioned study directed by Olweus (1993). From that moment on, several big scale research projects have been developed and implanted in the United States, the UK, Sweden, Japan, and Australia.

Koo (2007) depicts the history of bullying in the UK claiming that it has been well documented from the 19th century onwards and attributing its causes to different features of the victims. Furthermore, this aggressive behaviour can already be found in the Victorian novel by Hughes (1870) "Tom Brown's School Days: By an Old Boy": *"But after a sound night's rest the old boyspirit comes back again. East comes in reporting that the whole House is with him, and he forgets every thing except their old resolve, never to be beaten by that bully Flashman"*. This indicates that, even if bullying was not typified as an offense in England, it was already a recognised issue.

It is generally accepted that a concept such as bullying is strongly related to social culture. Therefore, it seems nothing but pertinent to explore the evolution of the concept in Japan, which is one of the few countries in Asia in which there has been research of the matter, at least until very recent times. Koo (2007) describes bullying in the 1970s and 1980s in Japan as one of the biggest social issues featured in Japanese mass media and points to the subtle differences in meaning between *ijime* (term in Japanese language) and bullying. Those differences are also appointed by Kanetsuna and Smith (2002), which explicitly compares it to the British concept of bullying, highlighting those regarding the number of students participating in the aggression, which were higher in Japan; the age difference between the attackers and the victims, more significant in the UK; and the fact the whereas bullying in Britain was more physical, *ijime* aggression were mostly of psychological nature.

In the last few decades, remarkable amount research has been taking place in the United States as well. Nansel et al. (2001) depicted bullying behaviours stating that although conflict and harassment can be considered as part of youth life and development to a certain extent, bullying presents a potentially more serious threat to healthy youth development. Therefore,

later research suggests a growing concern about the effects of bullying on students. Hendricks et al. (2012) state that bullying has increased in the United States over the past thirty years which can be linked to the increasing number of suicide and homicide cases in some areas. Furthermore, Safran and Safran (2008) speculate that bullies bully others in order to maintain the imbalance in power.

Second, there have been several studies that have pointed out that there is also a conceptual difference in the understanding of bullying, which means that bullying represents different implications for students, teachers, and parents.

Arora (1996) presents the results of a research in which students of secondary education (13/14 years old) had to give a definition of bullying. The study showed the frequency of the elements included in their definitions, being the most common: direct physical aggression, the attacks made from a group to a single person, the frequency of the attacks, making fun, swearing, or teasing. In addition, this article claims that the younger the students, the greater the tendency to relate bullying to physical aggression.

Mishna et al. (2005) carried out a survey in which teachers were presented with the aforementioned definition by Olweus and were asked to give a definition. The teachers highlighted the existence of power imbalance, intention, both direct and indirect aggressive behaviours. There was disagreement among the surveyed in the equality of importance of physical versus psychological attacks. Finally, most of the teachers did not consider repetition as one the main characteristics of bullying behavioural patterns.

In a similar study in which several different attitudes were presented to teachers and students in order for them to categorise them as bullying, Boulton (1997) compares teachers and students' answers and the results show that, with the exception of name calling, a significantly greater proportion of teachers than pupil viewed the relevant behaviour as bullying.

Third, at this point it becomes necessary to consider the concept of cyberbullying not only as a the strongest subcategory and materialization of bullying but also as a definition-changer due to the imminent implantation of social networks and technology in everyday human life.

Hinduja and Patchin (2014) define cyberbullying as *“willful and repeated harm inflicted through the use of computers, cell phones, and other electronic devices”*. They base this definitions in the constructs associated to the traditional definition of bullying: malicious intent, violence, and repetition. Imbalance in power is intentionally left out of the definition because the type of power exerted in cyberspace is *“somewhat amorphous and often shifting”*.

Hendricks et al. (2012) go through the evolution from bullying to cy-

berbullying highlighting the fact that while elementary school children are confined to physical and verbal bullying due to their limited access to technological devices, middle and high school students instigate bullying through their electronic devices.

Corcoran et al. (2015) underline the difficulty to define cyberbullying in a world which is in constant flux pointing out that by the time researchers reach consensus, young people will be using technology that did not exist when the consensus was achieved. Furthermore, they put into question if cyber aggression can be classified as bullying arguing that, the characteristics of the former can be in fact different to the ones of the latter. It is also speculated that a better approach to defining and measuring cyber-based aggression would support better intervention and prevention methods intended to reduce it.

Kleinknecht et al. (1997) define social anxiety as *“The experience of intense anxiety or fear associated with social situations in which one might be viewed or scrutinized by others.”* Pabian and Vandebosch (2016) relate social anxiety to bullying and cyberbullying. The results of their study indicated that social anxiety can be found as a predictor of victimization, rather than as consequence. Those suffering from social anxiety were more likely to be a victim of both bullying and cyberbullying and, those who perpetrate traditional bullying, but not cyberbullying, were related to subsequent social anxiety. However, social anxiety was not found to be a risk factor for the perpetration of (cyber)bullying.

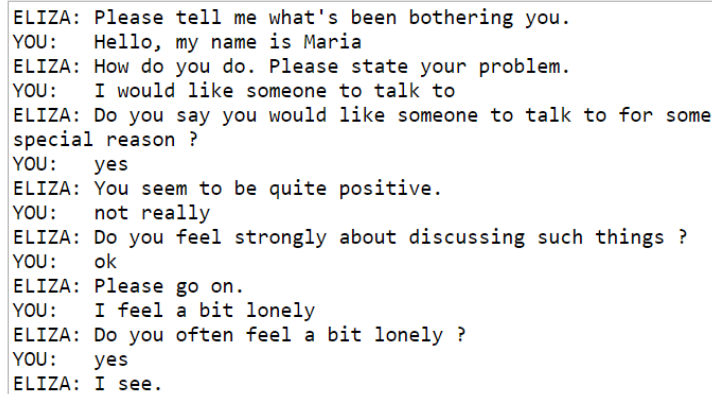
2.2. Chatbot technology as a tool for bullying

2.2.1. History of chatbots

ELIZA, the earliest chatbot, dates back to 1966. It was created at the MIT Artificial Intelligence laboratory and it simulates the behaviour of a psychotherapist. It simulates human conversation with the use of pattern matching and scripted responses. Although it creates the illusion of understanding, it has no built in framework for contextualizing events (Weizenbaum (1966)). Figure 2.2 shows an example of user interaction with ELIZA.

PARRY was created in 1972 by Kenneth Colby in the Stanford Artificial Intelligence laboratory. It simulated the behaviour of a patient with paranoid schizophrenia. It was tested in the early 1970s by a variation of the Turing Test in which a group of psychiatrists had to analyze a combination of real patients and computers running PARRY through teleprinters. The experts were able to identify them correctly 48% of the time. Colby regarded PARRY as a tool to study the nature of paranoia, and considered ELIZA

Eliza



```
ELIZA: Please tell me what's been bothering you.
YOU: Hello, my name is Maria
ELIZA: How do you do. Please state your problem.
YOU: I would like someone to talk to
ELIZA: Do you say you would like someone to talk to for some
special reason ?
YOU: yes
ELIZA: You seem to be quite positive.
YOU: not really
ELIZA: Do you feel strongly about discussing such things ?
YOU: ok
ELIZA: Please go on.
YOU: I feel a bit lonely
ELIZA: Do you often feel a bit lonely ?
YOU: yes
ELIZA: I see.
```

Figure 2.2: Example of conversation with ELIZA.

as a potential clinical agent who could, within a time-sharing framework, autonomously handle several hundred patients an hour (Colby (1999)). Figure 2.3 shows an example of conversation with the paranoid schizophrenic chatbot, PARRY.

Jabberwacky was developed in 1988 by Rollo Carpenter and went online in 1997. It was meant to simulate natural human chat in an interesting, entertaining and humorous manner. The general AI of Jabberwacky stores everything everyone has ever said, and finds the most appropriate thing to say using contextual pattern matching techniques (Carpenter (2011)). As can be appreciated in figure 2.4, Jabberwacky's responses are influenced by what other users have said to him before.

Dr.Sbaitso was commercialised in 1992 as an Artificial Intelligence speech synthesis program created for MS DOS-based personal computers. It simulated the behaviour of an stereotypical psychologist giving its answers not only in text but also in a mechanical voice produced by an internal synthesizer. Its answers were computed using pattern-matching techniques, similar to those used by ELIZA (Candello and Pinhanez (2016)). Figure 2.5 shows an example of conversation with the Dr.Sbaitso.

In 1995 Dr. Richards S. Wallace invented A.L.I.C.E, the Artificial Linguistic Internet Computer Entity. It was a natural language processing bot inspired as well by ELIZA. It applied heuristic pattern-matching rules to the

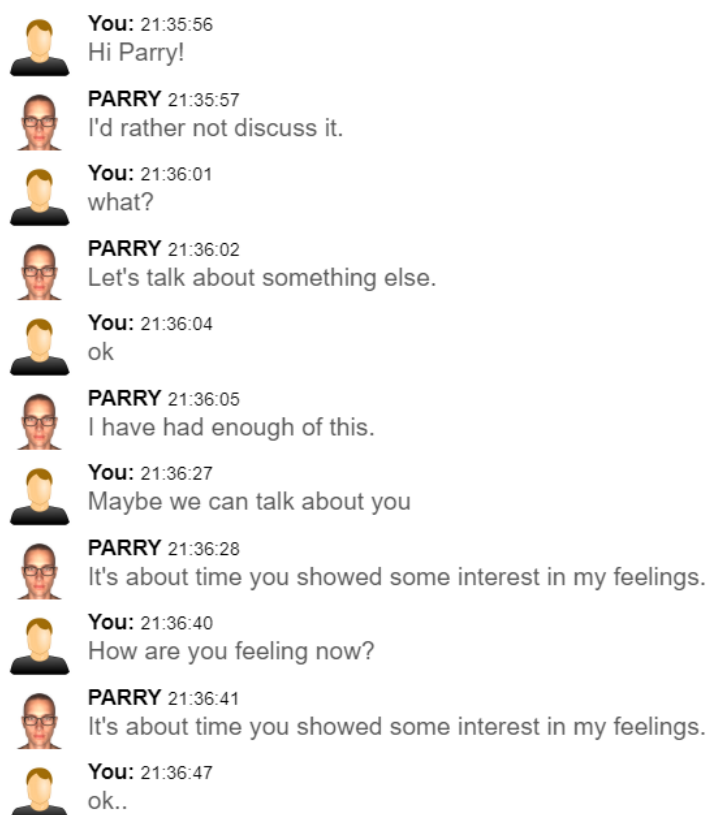


Figure 2.3: Example of conversation with PARRY.

user input in order to have conversations and was able to win the Loebner Prize (formal contest of a Turing test) three times (Candello and Pinhanes (2016)).

Six years later Smarterchild was developed by ActiveBuddy, Inc. as an intelligent agent widely distributed across global instant messaging and SMS networks. It was capable of giving information about the weather, stock, news, yellow pages listings, etc. More than nine million users interacted with SmarterChild in its first year, and it is considered to be a precursor to Siri and Voice by Apple and Samsung, respectively.

Mitsuku was created by Steve Worswick in 2005. It claims to be an 18-year-old female chatbot from Leeds, and it contains all of Alice's AIML files, with many additions from user generated conversations. Its intelligence includes the ability to reason with specific objects and it is a four-time Loebner Prize winner. As shown in figure 2.6, Mitsuku's responses have a more complex structure and a considerable relation to context.

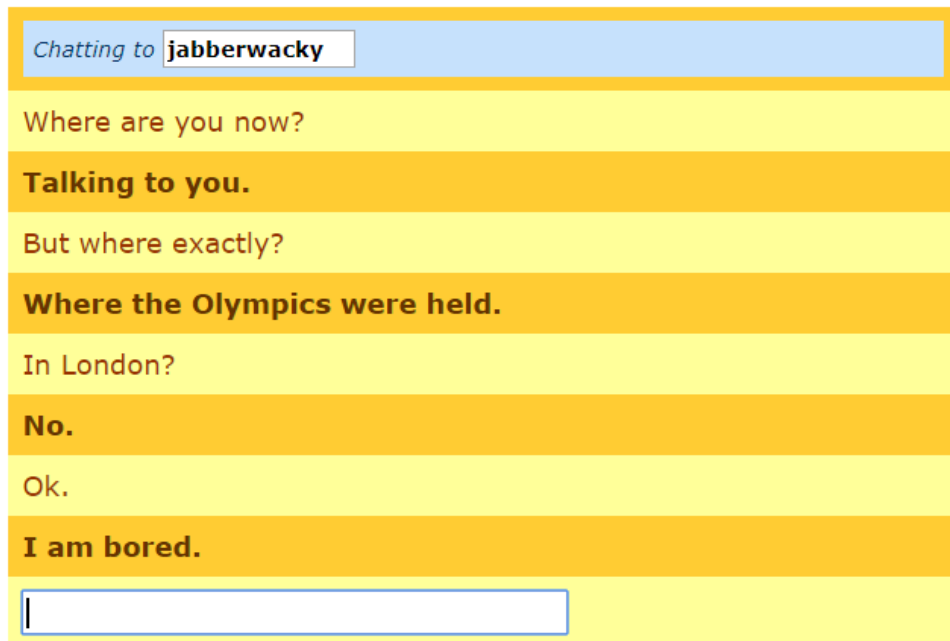


Figure 2.4: Example of conversation with Jabberwacky.

In 2006, IBM built Watson, a question answering computing system that was designed to apply advanced natural language processing, information retrieval, knowledge representation, automated reasoning, and machine learning technologies to the field of open domain question answering. Watson uses IBM's DeepQA software and the Apache Unstructured Information Management Architecture framework (UIMA). It runs on the SUSE Linux Enterprise Server 11 operating system using Apache Hadoop framework to provide distributed computing. It was the first bot to win the television show *Jeopardy!* (High (2012)).

Siri was created as an intelligent personal assistant by Apple in 2010. It uses Automatic speech recognition to translate human speech into text. Using natural language processing, it translates transcribed text from voice into parsed text. Its technology takes into account question and intent analysis (Deshpande et al. (2017)). Figure 2.7 shows Siri's transcription of a voice conversation with Siri.

Only 2 years after Siri, Google created Google Now, which employs natural language user interface to answer questions, make recommendations, and perform actions by passing the requests to a set of web services (Candello and Pinhanez (2016)).

Alexa was a voice service inhabiting the Amazon Echo device that was created in 2015. It uses natural language processing algorithms for voice in-

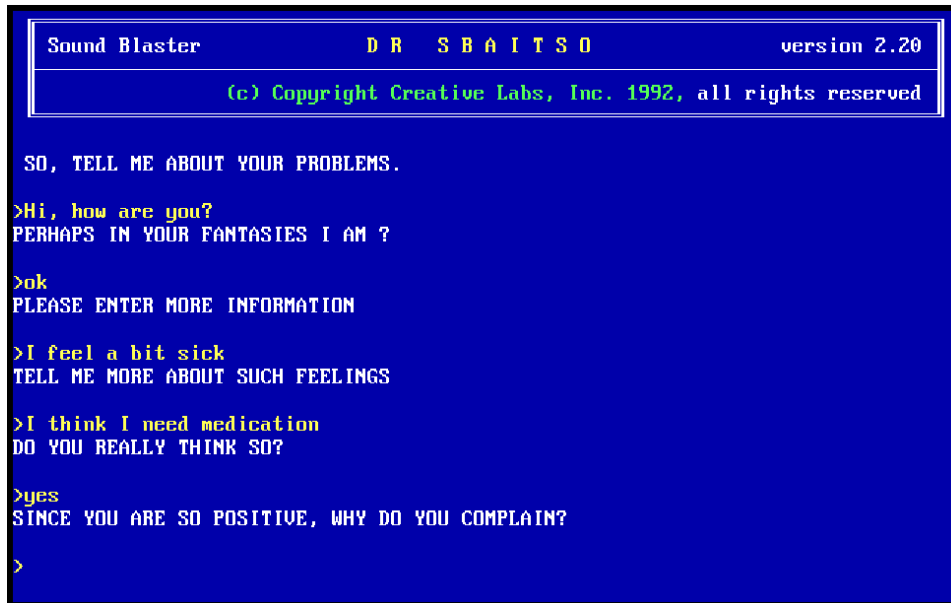


Figure 2.5: Example of conversation with Dr.Sbaitso.

teraction. It is capable of music playback, making to-do lists, setting alarms, streaming podcasts, playing audiobooks, and providing weather, traffic, and other real time information. Alexa can also control several smart devices using itself as a home automation hub (Deshpande et al. (2017)).

Cortana was first launched in 2014 for Windows Phone 8.1, inspired by the character in the Halo videogame series. Cortana has a Notebook which stores personal information preapproved by users. It also learns by user's phone usage, location and communication. It is activated by voice and text input and tries to understand the basic context of a conversation. Cortana relies on Bing's backend services and is backed up by thousands of servers in the background (Candello and Pinhanez (2016)).

Another group of conversational system is increasingly becoming available, characterised by direct integration to smartphone messenger systems like Facebook or Telegram. In 2016, Facebook launched the Messenger platform service which allows developers to create chatbots which can interact with Facebook users.

In 2016 Microsoft launched a chatbot on Twitter, known as Tay. Tay was designed to mimic the language patterns of a 19-year-old American girl, and to learn from interacting with human users of Twitter. This caused people to interact with the bot inappropriately, using offensive language and content which ultimately caused Microsoft to take the chatbot down after only 16 hours (Cameron et al. (2017)).

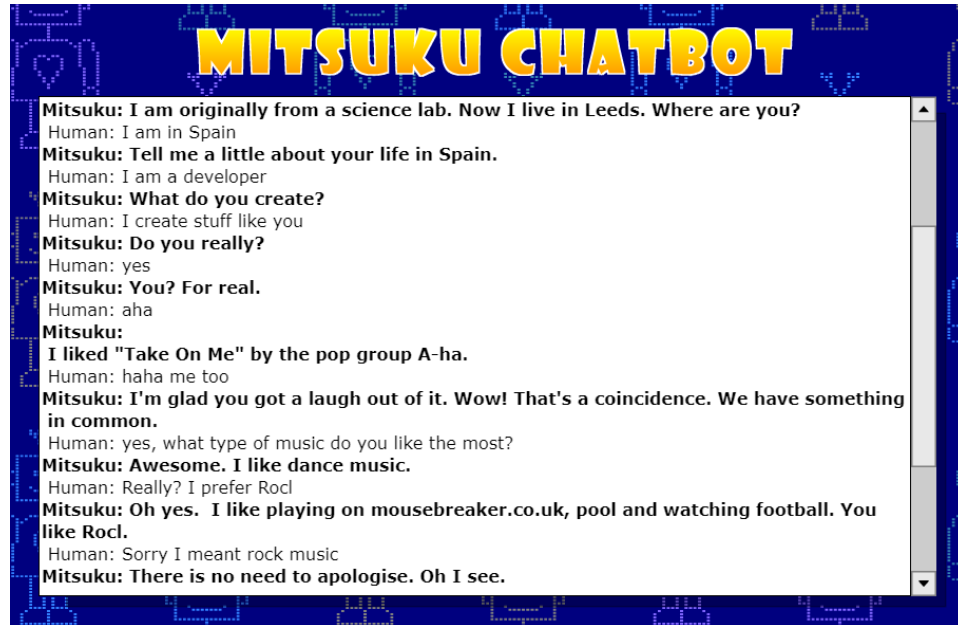


Figure 2.6: Example of conversation with Mitsuku.

2.2.2. Different chatbot configurations

There are several different approaches that can be used to develop a chatbot. According to Hussain et al. (2019), these approaches can be divided into three main categories:

- **Rule-Based Approaches:** They are the traditional methods, which have been widely used in most commercial implementations and often adopt hand-crafted rules to select the most likely result. These systems are prone to frequent errors as the most likely result is not always the desired one.
- **Retrieval-Based Approaches:** They choose a response from candidate responses. The key to retrieval-based methods is message-response matching. Matching algorithms have to overcome semantic gaps between messages and responses (Chen et al. (2017)).
- **Generative-Based Approaches:** They generate proper responses during the conversation and can generate more proper responses that could have never appeared in a human-human conversation corpus. Generative models overcome this reliance on predefined responses by generating new responses that are built from applying a set of techniques. These models are usually based on various machine translation techniques (Ramesh et al. (2017)).

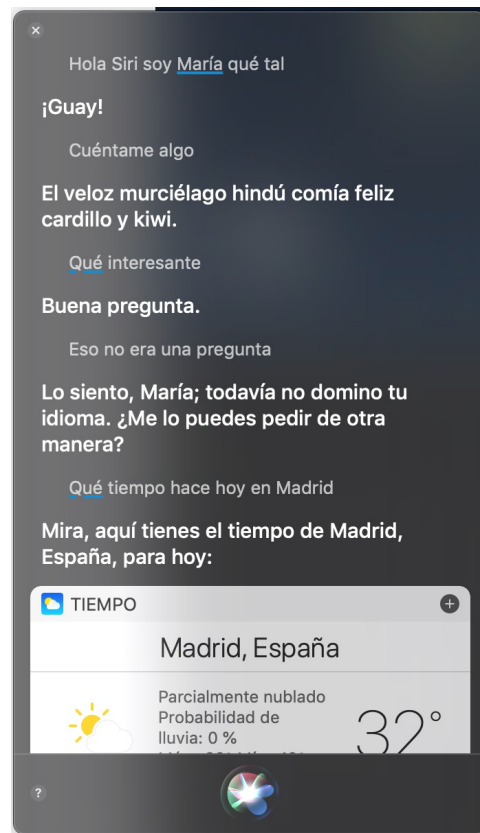


Figure 2.7: Example of conversation with Siri.

At the same time, these main approaches can employ a combination of the following techniques:

- **Parsing:** this method consists on taking the text as input and extracting meaningful information that can be categorized, stored, and manipulated. ELIZA used parsing in order to extract the keyword in a sentence, match it against the corpus documents, and find an appropriate answer (**lexical parsing**). Modern chatbot systems like Dialogflow use another type of parsing which converts the input to a machine-understandable representation of its meaning (**semantic parsing**).
- **Pattern Matching:** it is widely used in question and answer chatbots. This technique classifies the input as *patterns* in order to produce a suitable response stored as a *template*. The patterns and templates have to be handcrafted by the developer, which makes pattern matching a limited approach.

- **AIML:** Artificial Intelligence Mark-up Language takes advantage of keywords in sentences, and generates a predetermined output. By verifying the keywords, and the ways it could be used, the chatbot would be able to generate an output that the user needs. AIML contains data objects consisting of two types of units called topics and categories. On the one hand, the topic contains the name, attribute and a set of categories related to the topic. On the other hand, category is used to represent the knowledge gained through the input and contains a template which represents the chatbot response (Teckchandani et al. (2018)).
- **Chatscript:** It aims to be easier to maintain than AIML by focusing on better syntax, it fixes the zero-word matching problems. The Chatscript first finds the best topic that matches the user query string and executes a rule in that topic. Rather than using separated categories for each word as in AIML, Chatscript uses *concepts* to merge similar words with meanings or parts of speech (AlHumoud et al. (2018)).
- **Ontologies:** It is a semantic network, a set of hierarchically and relationally interconnected concepts. These concepts can have natural language names and can be used directly in chatbots, to figure out hyponyms, synonyms and other relations between the concepts. The advantage of ontologies is that the concepts are interconnected into a graph, which enables computers to search through and, using special reasoning rules, imply new statements (Bradeško and Mladenović (2012)).
- **Markov Chain Model:** It is used in Chatbots to build responses that are more applicable probabilistically and, consequently, are more correct. The idea of Markov Chains is that there is a fixed probability of occurrences for each letter or word in the same textual data set (Bradeško and Mladenović (2012)).
- **Artificial Neural Networks Models:** Deep learning aims to imitate the human brain function in terms of data-processing and pattern-creation that are used for decision making. Many different variants of artificial neural networks are used for natural language processing and conversational modelling, such as recurrent neural network, sequence to sequence, and long short term memory networks (Hussain et al. (2019)).

2.2.3. Previous work on bullying using chatbot technology

As described in previous sections, bullying and cyberbullying are issues that affect the world's population, especially children and teenagers. Ac-

According to Zych et al. (2017), in Spain bullying perpetration rate is of 5%, bullying victimization rate is 23.2%, bully/victim status rate is 16.4%, cyberbullying perpetration rate is 4.3%, cyberbullying victimization rate is 13.1%, and cyberbully/cyber victimization rate is 10.8%. For this reason, it is common to find studies that try to confront and offer significant solutions to the issues of bullying and cyberbullying.

Latham et al. (2010) was one of the earliest approaches to the support of bullying in large organizations through conversational expert systems. The information about harassment and bullying policies was represented using a knowledge tree combined with a graphical interface which enabled traversing the tree. The result of this development was Adam, a conversational agent capable of giving simple and direct answers to natural language questions. An independent evaluation was conducted in order to explore Adam's effects and the results showed that the system allowed users to access information about sensitive issues like bullying anonymously and at any time, aiding organizations in positively implementing harassment policies and improving workforce cohesion.

Schussler et al. (2017) have created Eli, a potentially bullied chatbot in order to enhance teacher candidates' skills in responding to bullying. In the study, an experiment in order to evaluate the effectiveness of the project was conducted. The results showed that the candidates that used Eli in their training, had a more refined word choice, maintained better eye contact, and did not rely on their notes during regular role-play, which suggests a greater comfort level than those whose practice consisted of reflection.

A very interesting study was the one conducted by Lee et al. (2019), which created a self-compassion chatbot called Vincent in order to assess how people can care for another being as a way to care for themselves. The experiment compared self-compassion scores before and after two weeks of interaction with Vincent. The chatbot was built with Google's Dialog flow and was integrated to Facebook Messenger. In order to have better comparability among participants, they used limited pre-set responses from which participants could choose. Vincent was given caregiving and care-receiving conditions and the results showed that while self-compassion increased for both conditions, only those with care-receiving Vincent significantly improved.

Another significant contribution is the one made by Young Oh et al. (2019), which explores the effects of conversational bots in relation to the Students' attitude change towards bullying. For the development of the study, fifth-grade students chatted to one of three different conversational agents: a bully bot (1), a bullied bot (2), and a teacher bot (3). The aim of the study was not only to explore the effects of an anti-bullying program that utilizes a chatbot on attitude change toward bullying problems, but also to compare the results of the chatbots' different roles among groups. The

results showed, on the one hand, that conversation with the virtual agent was effective to positively change students' attitudes toward bullying problems and, on the other hand, that the agent's different role had also an effect on the students' anti-bullying attitude change.

2.3. Other technological approaches to bullying

The previous section explored the most relevant studies which use chatbots as a tool for issues derived from bullying. However, since bullying is a problem worrying most of the world population nowadays, chatbots are not the only valid solution proposed by researches along the globe. Technological development has given a wide range of possibilities to confront the bullying issue: from anti-bullying mobile apps to automated detection of cyberbullying in online social networks.

A relevant example of the latter is the study by Zhong et al. (2016), which explores the relationships between text and visual content with respect to cyberbullying. The aim is to find a correlation between shared media in the form of posted images and captions, and the occurrence of cyberbullying events, with an eye on the development of early-warning mechanisms. For the cyberbullying detection, neural network technology was used and, with regard to early-warning mechanisms, the probabilities of creating images being triggers of cyberbullying events were calculated. The results showed that, in the context of photo-sharing, captions in particular can serve as a surprisingly powerful predictor of future cyberbullying for a given image.

In the study by Hall et al. (2004) a Virtual Learning Environment populated by synthetic characters for children to explore issues surrounding bullying behaviour was presented. The results revealed that children expressed the most favourable views towards the characters and the highest levels of believability towards the bullying storyline.

SMART Talk is a program designed to teach adolescents how to solve interpersonal problems without violence Bosworth et al. (2000). The program tries to address three major components in order to manage the students' anger and resolution. The first component is anger management and uses animation, interactive assessment interviews, and games to teach students to recognise the cycle of anger. In the second component, perspective talking, students have to view anger situations from the perspectives of all participants in the situation, through a game format. The final component is dispute resolution and it includes an interactive mediation process that guides students to generate solutions to their conflict. The results indicated that the intervention was successful in diminishing students' beliefs supporting violence and increasing their intentions to use nonviolent strategies.

Chapter 3

Architecture and Implementation

For the development of this project, a fully integrated architecture was designed, developed and tested. In this chapter, the main components forming the system will be depicted in detail, giving both technical and functional documentation of the development process.

3.1. Overview

The architecture of OttoBot has been built up using different modules that address the different necessities of an online chatbot architecture. As depicted in figure 3.1, the main modules of OttoBot are interconnected in order to conform the system. The modules have been arranged in two main groups: the user interaction modules and the internal dialog-management modules. The module in charge of presenting the system to the final user as well as connecting it to the engine of the bot is called user interaction and will be explained in section 3.2. On the other hand, the user message and language processing, the context management, the data storage and analytics, and the bot engine deployed as a UNIX service are thoroughly described in section 3.3. The whole project of OttoBot can be found in GitHub: <https://github.com/mariaesp/OttoBot>.

3.2. User interaction modules

In this section the modules in charge of connecting OttoBot to the final user will be defined. The main component used in these modules is the technology offered by Google Hangouts API and the different choices that were made in order to connect it to OttoBot's backend.

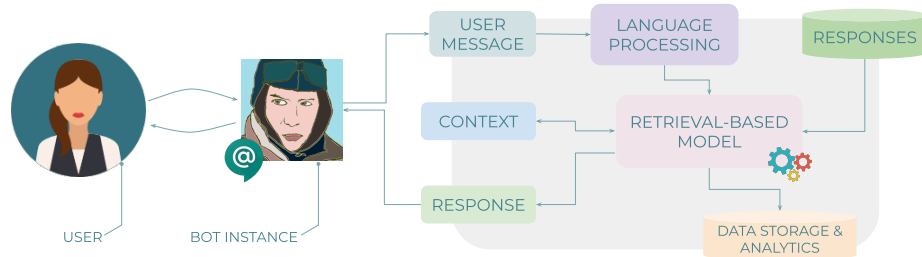


Figure 3.1: Modules that conform the system's main architecture.

3.2.1. User Interaction

The architecture includes a module dedicated to user interaction, which is in charge of connecting the user and the bot in order for them to communicate. The user interacts directly with OttoBot through Google Hangouts Chat, the main chat application from Google, as it can be seen in Figure 3.1. In the development process, several chat applications were considered: WhatsApp, Facebook Messenger, Telegram, and Google Hangouts.

Initially, WhatsApp Messenger was the preferred option because it is the most widely used in Spain, as depicted in Figure 3.2 (Moreno (2018)). Despite being the most popular, WhatsApp was finally discarded because the use of the WhatsApp API for automated responses, as is the case of a chatbot, is forbidden in their terms and conditions and it results in a ban of the phone number.

Having in mind this limitation, the next popular option would be Facebook Messenger. Messenger supports the creation of conversational agents and there are several platforms that offer free support for their development and implementation free of charge. The main issue regarding this option was that it requires the users to be registered in Facebook and also to have the Facebook Messenger app installed in their devices, which would make more difficult for users to try it out.

Another option that was considered was Google Hangouts Chat. Although it is not the most widely used in Spain, the only requirement is to have a Google Account which, at least every Android user has. It is important to notice that in Spain, 90% of smart phone users have an Android device and have, therefore, a Google account. Figure 3.3 shows the result of the market share for operating systems in Spain at the end of 2018, as revealed by a study made by Kantar Worldpanel (2019). In addition, Google apps come pre-installed in every Android device, which makes it easier for the user to reach the chatbot. Furthermore, Google also supports chatbot

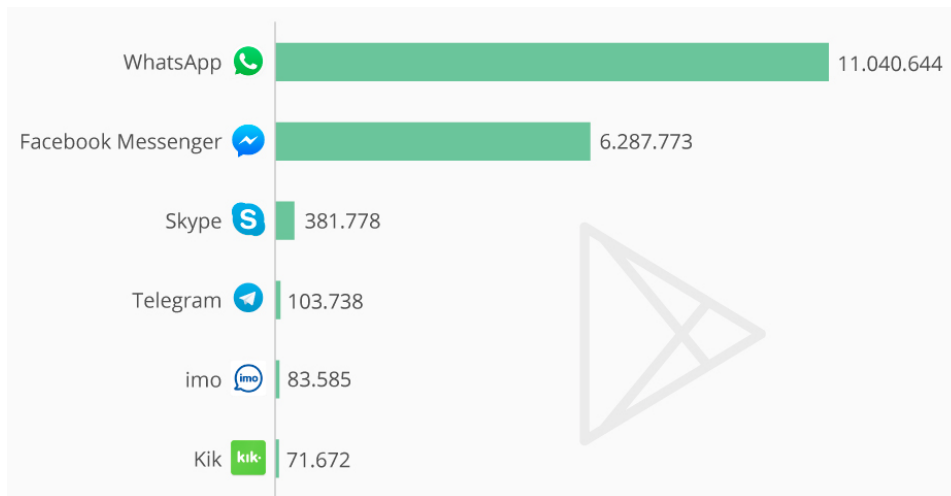


Figure 3.2: Number of daily active users in the different instant messaging apps in Google Play in Spain (April 2018).

development for Google Hangouts and offers an API and several tools in order to develop and implement a conversational agent. It is important to notice that Google allows all the users under the Complutense University of Madrid domain to use all their developer tools for free, which make the development of OttoBot more complete.

Bearing in mind the previously exposed facts, added to the fact that all the students, professors, and administrative staff of the Complutense University of Madrid also have an institutional Google Account, Google Hangouts Chat was chosen as the conversational agent chat platform.

In order to talk to the bot, the only requirement is to sign in with a Google account, go to Google Hangouts Chat, search for OttoBot and start chatting with him. As shown in Figure 3.4, the interaction with OttoBot is like any other chat in Google Chat and works both from the smart phone app and the computer. OttoBot is able to reply whenever the user sends a message, but it will not be the first one to write to users. Therefore, the user is always in charge of deciding when is right time to talk to OttoBot.

It is also important to notice that one of the main advantages of a modular architecture is that the chat client can easily be changed without altering the rest of the application. As it will be explained in Section 3.2.2, the only part that would need to be adapted is the communication with the API which, instead of sending requests and responses to Google Hangouts Chat API, would have to send them to the new selected client, like Facebook, Telegram, or even WhatsApp.

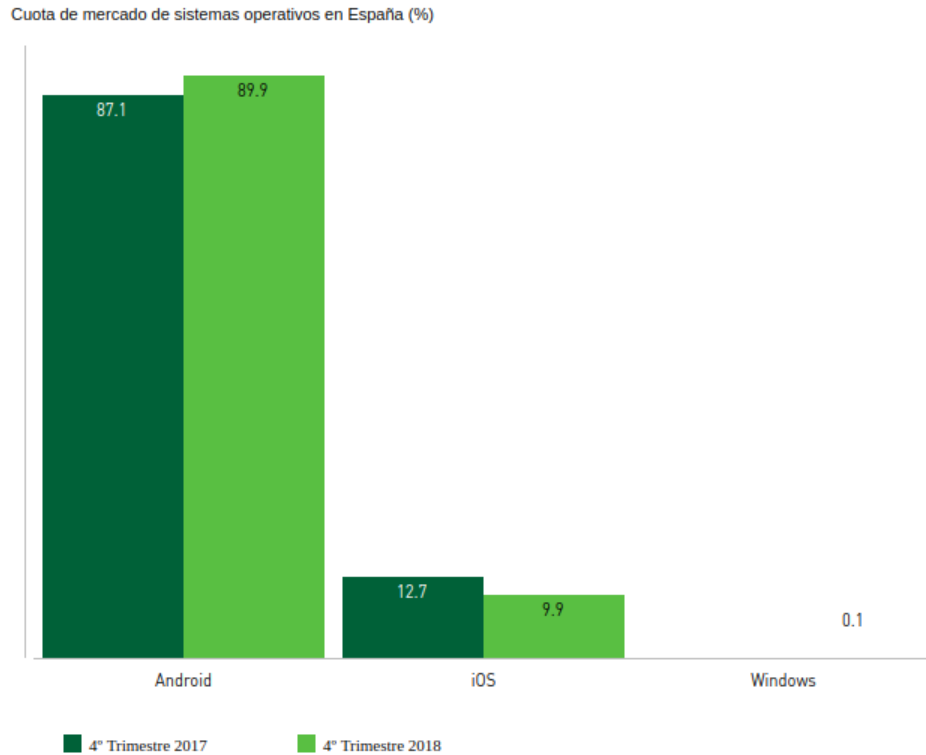


Figure 3.3: Market share for operating systems in Spain.

3.2.2. Google Hangouts Chat API

As explained in Section 3.2.1, OttoBot was developed using Hangouts Chat API, one of the products offered by Google Developers. In the documentation for Hangouts Chat API, the user can find information such as how to create bots right into the Hangouts Chat, letting users get information, take quick action, and stay notified without leaving the conversation.

The endpoint type chosen in order to deliver events from Hangouts Chat to OttoBot app was HTTP, that is, presenting OttoBot as a web service. Communication between client computers and web servers is done by sending HTTP Requests and receiving HTTP Responses. The details on the server deployment will be explained later, in section 3.3.4.

In order for the bot to start receiving and sending messages, a configuration process must take place. First of all, the bot must be registered specifying the URL in which the bot is hosted in the Chat API configuration, as well as some other options such as the profile picture, the scope of the bot (single or group conversations), who has right to contact the bot, etc.



Figure 3.4: Example of simple conversation with OttoBot.

Once the bot is registered and published, Hangouts Chat is able to recognise events addressed to the bot and dispatch them to the correct endpoint.

In addition to the above-described process, Google includes a security layer in order to make sure the requests processed by the bot application are actually coming from Google. Hangouts Chat includes a bearer token in the Authorization header of every HTTP Request to a bot.

```
POST
Host: yourboturl.com
Authorization: Bearer AbCdEf123456
Content-Type: application/json
User-Agent: Google-Dynamite
```

The string `AbCdEf123456` is the cryptographic token produced by Google that should be checked in the receiving end.

When the bot receives an event from Hangouts Chat, the body of the HTTP request includes all the event-related information such as the type,

the time, the message, etc. The following JSON is an example of an event containing the message *hola* received by OttoBot:

```
{
  type: 'MESSAGE',
  eventTime: '2019-08-25T17:24:43.485039Z',
  message: {
    name: 'spaces/iKq64gAAAAE/messages/yyf3kHY_
    l-Q.yyf3kHY_l-Q',
    sender: {
      name: 'users/113233486793214437597',
      displayName: 'MARIA SOLEDAD ESPINOSA RUIZ',
      avatarUrl: 'https://lh3.googleusercontent.com/
      a-/AAuE7mAukXgIdrMmUkDuTjzUiUair6n67T-V4SQi3Xm3',
      email: 'mariases@ucm.es',
      type: 'HUMAN'
    },
    createTime: '2019-08-25T17:24:43.485039Z',
    text: 'hola',
    thread: {
      name: 'spaces/iKq64gAAAAE/threads/yyf3kHY_l-Q',
      retentionSettings: [Object]
    },
    space: {
      name: 'spaces/iKq64gAAAAE',
      type: 'DM'
    },
    argumentText: 'hola'
  },
  user: {
    name: 'users/113233486793214437597',
    displayName: 'MARIA SOLEDAD ESPINOSA RUIZ',
    avatarUrl: 'https://lh3.googleusercontent.com/a-
    /AAuE7mAukXgIdrMmUkDuTjzUiUair6n67T-V4SQi3Xm3',
    email: 'mariases@ucm.es',
    type: 'HUMAN'
  },
  space: {
    name: 'spaces/iKq64gAAAAE',
    type: 'DM'
  },
  configCompleteRedirectUrl: 'https://chat.google.com/api/
  bot_config_complete?token=AC1qJQTklc87oH__p_4Hp_5eo0FA
  sj117SyGC35TeYaXIe5hQPXud7Fj8fzBrcq9prvK8x0ifXP7qhpCWs0
```

```
vdGoIOSSrnIJzKbTrBrBSeFxE2bNomz8ufzmNE5q_b-Q_dSmHHtBG  
XMqXRk9Pkd6'  
}
```

The bot responds to events in a synchronous way, which means that it returns a JSON-formatted message payload in the HTTP response within 30 seconds of having received an HTTP request. More information on how the HTTP response is sent and how responses are tailored will be given in the following sections.

3.3. Internal dialog-management modules

This section will describe the main functionality of the system, namely the development of all the bot logic in JavaScript, from message processing to response tailoring, the storage of a history of conversations for further analytics, and the server in which the application is running. The main module corresponds to all the components over the gray area in Figure 3.1.

The code covering the functionality of the following subsections is written in JavaScript and executed with Node.js (version 12.9.0), which is a JavaScript runtime built on Chrome's V8 JavaScript engine.

3.3.1. User Message and Language Processing

As explained in section 3.2.2, the messages written to OttoBot by the users are sent as HTTP requests from Google Hangouts Chat to OttoBot application. In order to receive these requests, the application makes use of Express, a Node.js web application framework that provides a robust set of features for web and mobile applications, such as HTTP utility methods which can help creating a robust API. In the case of OttoBot, once a secure connection is established using Google Authentication and Google Keys, the Express methods `listen()` and `post()` are used in order to receive and send messages.

Once a user message is received, the application takes the text content written by the user from the HTTP request and applies a series of transformations in order to make it comply to the retrieval-based model and the resulting text is sent to the module explained in next section in order to retrieve a meaningful response.

The transformation process as well as the model will be extensively explained in Chapter 4.

3.3.2. Context, Retrieval-Based Model, and Responses

The relation of the responses given by the conversational agent to the context of the conversation has always been a critical issue in chatbot development. The more context-related are the answers given by the bot, the more human-like they will be.

In order to make OttoBot’s responses more human-like, a context has been constructed as part of the main application. Context construction includes changes in the bot’s state derived from every interaction it has with the user. Since this bot is meant to be chatting with psychologists, three variables related to the bot’s trust in the professional, its acceptance of the problem and its commitment to take action about it have been added to the context. Chapter 4 will describe these variables and their effects on the responses of the bot.

Taking into account the context and the user message, the retrieval-based system is queried for a response. This system uses AIML 2.0, which is the second release of the standard XML specification designed for scripting chat bot responses in the late 90s. AIML 2.0 also has a tag-based representation, which makes it possible to capture the basic structure of a pattern path and a hierarchical response template.

The file containing all the possible request-response mappings that conform OttoBot’s *brain* is written in AIML, which has the following form:

```
<aiml>
  <category>
    <pattern>HOLA</pattern>
    <template>Hola, ¿Qué tal?</template>
  </category>
</aiml>
```

AIML defines how a bot should respond to a question using `<pattern>` and `<template>` elements. A pattern represents the user’s question, and the template defines the bot’s response. The pattern and template elements are grouped in a category element, which can be grouped under topic elements. Therefore, when the retrieval system is queried with the text “*hola*”, it will reply “*Hola, ¿qué tal?*”.

There are more than 50 different tags included in the AIML documentation, Table 3.1 provides a summary of them and description of the AIML tags used for the development of OttoBot.

In order to connect the AIML file with the rest of the application, an AIML parser for JavaScript has been used. The parser, AIMLInterpreter (TerribleNews (2016)), is a module that allows the user to parse AIML files

Tag	Description:
<code><aiml></code>	The AIML root element delimits a block of AIML code. All other elements must be descendents of the root element
<code><bot/></code>	The bot element is used to recall custom bot properties defined in the .properties file. These variables are accessible to all users of the bot
<code><category></code>	The category element delimits a base unit of knowledge in an AIML-based chatbot. In a very broad sense, a single category accepts an input, and returns an output
<code><condition></code>	The condition element is used to create an IF-THEN-ELSE type of control flow within a bot's response. This is done by checking the value of a predicate, and returning a response depending on that value
<code><get/></code>	The get element is used to return the value of a particular predicate
<code></code>	The list item element can be a child of both <code><condition></code> and <code><random></code> . It makes it possible to attach multiple responses, each of which is chosen under certain circumstances
<code><pattern></code>	The pattern element is the block within each category that defines a linguistic pattern against which the user's input can be matched
<code><random></code>	The random element can be used in conjunction with list item elements to provide a set of potential bot responses, one of which will be returned at random in the case that the category is matched
<code><set></code>	The set element is used to set a predicate variable. Predicates are not hardcoded like properties, and can be initialised during conversation
<code><srai></code>	The srai element allows the bot to recursively call categories after transforming the user's input
<code><star/></code>	The star element is used to echo portions of the user's input that were captured by wildcards
<code><template></code>	The template element is the portion of category that defines its return value. It can contain a variety of other elements. Every category must contain a template block
<code><that></code>	The that element is an optional child of the category element that is used to establish the context of the pattern. If a category contains a that element, the pattern can only be matched if the last sentence of the bot's previous response matches the contents of that
<code><think></code>	The think element allows the bot to set predicates without actually displaying the contents of a set element to the user
<code><topic></code>	The topic element allows the bot to contextualise categories according to the value of a predicate named is "topic"

Table 3.1: AIML tags description table.

and to find the correct answer to a given message. The code is mostly based on the reference implementation for AIML2.0, Program AB, which was written in Java. Although the project also contains several AIML files and configurations in order to enhance user's bots, they are only valid for English speaking agents. Therefore, only the parser part of the project was used for the development of OttoBot.

AIMLInterpreter allows the use of most AIML 2.0 functionalities such as the use of bot properties, and the conditional tags. These have been used in order to maintain the aforementioned context. The three variables related to the state of the bot are updated depending on the different possible interactions that the user and the bot can have. The maintenance of the context is done by updating those variables, which can be done using the `<condition>`, `<set>`, `<think>`, and `<srai>` tags in the following way:

```
<category>
  <pattern>INCREASE TRUST</pattern>
  <template>
    <condition name="trust">
      <li value="none"><think><set name="trust">low</set></think></li>
      <li value="low"><think><set name="trust">medium</set></think></li>
      <li value="medium"><think><set name="trust">high</set></think></li>
      <li value="high"><think><set name="trust">total</set></think></li>
    </condition>
  </template>
</category>

<category>
  <pattern>COMO ESTAS</pattern>
  <template>
    <srai>INCREASE TRUST</srai>
    bien, ¿y tú?
  </template>
</category>
```

In the example above, the first category is in charge of increasing the variable called *trust* by one level, depending on its current value. In order to use it, a category like the second one is needed. The `<srai>` tag allows to call that category whenever the bot receives the as input “¿Cómo estás?”.

Moreover, the `<condition>` tag allows the choice of different answers to a given input depending on the actual state of the bot in the following way:

```
<category>
  <pattern>QUIERES HABLAR</pattern>
  <template>
    <condition name="trust">
```



```
<li value="none">no</li>
<li value="low">no me apetece</li>
<li value="medium">Creo que sí</li>
<li value="high">Creo que necesito tu ayuda</li>
<li value="total">Me encantaría</li>
</condition>
</template>
</category>
```

This example shows the effect of the value of the variable *trust* on the responses given by the bot: depending on the *trust* level, the answer to “¿quieres hablar?” will range from no to yes.

Detailed information regarding variable choice and which interactions make their values increase and decrease can be found in Chapter 4, where the whole model for OttoBot is explained.

3.3.3. Data Storage and Analytics

The last piece of the architecture depicted in Figure 3.1 corresponds to the data storage. After every interaction with OttoBot a transcript of the user message and the bot’s response is recorded to a file. In addition, the values of the variables conforming the bot’s state are added to the file after each of the interactions.

Each user has a unique id number which is sent in every HTTP request. That id number is also used as the file identifier for each user, making it possible to store all conversations by user, even when conversations occur on different dates.

This functionality creates a history of interactions that can be used for several purposes in order to improve OttoBot’s behaviour such as machine learning.

For simplicity reasons, it was decided that the use of a database for this module would not be necessary. In addition, having each user conversation separated in its own file, makes it possible to send the users their conversations in a popular format for them to review their chat. However, the modularity of the project architecture makes it possible to add database storage, such as MongoDB, which is a general purpose, document-based, distributed database, just by adding the necessary code to the right module.

3.3.4. Bot engine as a UNIX service with systemd

The main module described previously in this section is executing in an Ubuntu Server which was provided by NIL group from the Software En-

gineering and Artificial Intelligence Department of the Computer Science Faculty of the Complutense University of Madrid.

The server has been accessed through using a SSH connection, which has made possible to deploy the application on the server. The Secure Shell (SSH) protocol is often used for remote terminal connections, allowing the user to access a text-mode terminal on a remote computer. In order to make OttoBot always available, even when the SSH connection was closed or when execution errors occurred, the application has been executed as a system daemon. A daemon is a service process that runs in the background and supervises the system or provides functionality to other processes. Traditionally, daemons are implemented following a scheme originating in **SysV** Unix. However, modern daemons should follow a simpler, more powerful scheme, as implemented by **systemd**.

In order to develop a service in Ubuntu, a file with a `.service` extension must be created in the `/etc/systemd/system` directory. In addition, the contents of the file must follow the following structure:

```
[Unit]
Description=OttoBot
After=network-online.target

[Service]
WorkingDirectory=/home/maria/OttoBot/
ExecStart=/usr/bin/node /home/maria/OttoBot/app.js
Restart=on-failure
Environment=PATH=/usr/bin:/usr/local/bin
Environment=NODE_ENV=production

[Install]
WantedBy=multi-user.target
```

Furthermore, in order to have OttoBot working in the Google Hangouts Chat API, it is necessary to use a public url using the HTTPS protocol. Therefore, another system service that binds the local 8181 port of the server to a public HTTPS URL using ngrok has been created. ngrok allows the exposure of a web server to the Internet.

The daemons described before can be easily controlled with the **service** command:

start a service	<code>sudo service ottobot stop</code>
stop a service	<code>sudo service ottobot start</code>
restart a service	<code>sudo service ottobot restart</code>
check the status of a service	<code>sudo service ottobot status</code>

In this Chapter, the architecture and all the related technical aspects have been described. As it has been previously stated, another important part of the development of the conversational agent is the model. In the next section, the conversational model of OttoBot will be explained in detail and, when needed, references to the concrete parts of the architecture will be given.

Chapter 4

A model of the bot's behaviour

In this chapter, a detailed description of the model will be given. The model of the bot's behaviour has been designed based on the core purpose of the bot's development, i.e. it must simulate the behaviour of a bullied teenager. In general terms, there are two types of chatbot responses: the ones generated from scratch as per machine learning models or the ones selected from a collection of possible responses previously established. OttoBot's model follows the latter due to the reliability and accuracy of the responses given which are, in a way, completely controlled by the developer.

The processing chosen for the user input has been reduced to the minimum possible for the maximization of the number of responses understood by the bot. Section 4.1 will detail the transformations done to the user input in order to achieve this purpose.

The response generation mechanism of OttoBot, as described in Chapter 3, is generated using AIML, a popular markup language used to define a set of rules with predefined patterns that act as the condition for the response generation. In addition, the rule-pattern system has been complemented with a context memory. That is, a representation of the state of the bot whose intention is to make OttoBot's responses more consistent with the rest of the conversation. The aforementioned context is composed by three state variables which are meant to represent the bot's state towards its relation with both the therapist and its victim condition: trust, acceptance, and commitment. It seems evident that this is a very limited model of the emotions that a person might be feeling during a therapy session. However, the objective was to create an approximation to that set of emotions with a bearable number of variables.

The final section of this chapter will describe the methodology developed in order to define the rules that conform the set of possible requests that

the bot is able to process. The aim of developing such methodology was to broaden that set and therefore make OttoBot more responsive to different kinds of inputs.

4.1. Adaptation of user input

One of the main limitations of AIML is that the input text must completely match the rule pattern. This means that even the use of upper case letters in a pattern that has been written in lower case will provoke that the said rule will not be triggered. For this reason, all the rules have been written in capital letters and the user's input is always transformed to capitals letters before reaching OttoBot's response system.

In addition, there are also symbols that have been removed from the input, such as the opening interrogation mark (*¿*), the opening exclamation mark (*¡*), and the accent marks (*á, é, í, ó, ú*). The opening symbols are used in Spanish language at the beginning of questions and exclamations. The accent mark is used in the vowels of words in Spanish to indicate where the stress of a word is. However, it is very common nowadays, especially in computer or mobile phone writing, to forget their addition to the sentence. For this reason, it was decided to eliminate them from both the rules and the input.

4.2. A computational model of a bullied subject's state

As it has been mentioned throughout Chapter 3, one crucial aspect of the bot's correct operation is the need for a context that would condition its responses. The term context here refers to the connection of a person to his or her relationships. Although a person is a unique entity, his or her reactions are directly related to the meaning derived from their social context, among others. In the case of OttoBot, it is an attempt to model a person's dialogue and that is the justification for the need of such a context. Thus the word context applied to OttoBot is used to define the relational involvement of the bot with people.

4.2.1. Variables representing subject-therapist relationship

On account of simplicity and effectiveness, the context construction has been reduced to 3 variables: trust, acceptance and commitment. In order to establish these variables, extensive research on therapy methodologies and approaches has been conducted.

Trust presents itself as a very straightforward variable. It is evident that in the development of any social relationship trust is a decisive element. Rotter (1967) is best known for its work in what regards trust. He defines it as the expectancy held by an individual or a group that the word of other(s) can be relied upon.

In relation to therapy, trust is also a fundamental characteristic. As Boszormenyi-Nagy and Krasner (1980) state in their study on trust-based therapy, *“trust building is both the ideological foundation and the primary tool of contextual therapy”*. In the study, they also point out the need of the therapist to present him or herself as worthy of the client's trust due to the fact that over time the client is meant to trust him with private aspects of his or her lives. Moreover, they establish four general ways in which trust between patient and therapist flows: (1) From the outset the therapist receives trust through the client's act of contracting for help, (2) the therapist has his own choice of how he or she will use their trust, (3) trust can be channeled into a one-to-one relationship between therapist and client, (4) rechanneling trust into the process of strengthening relationships between family members.

The other two variables are acceptance and commitment. Their election was based on research done regarding Acceptance and Commitment Therapy (ACT). *“ACT is a functional contextual intervention approach, which views human suffering as originating in psychological inflexibility fostered by cognitive fusion and experiential avoidance. In the context of a therapeutic relationship, ACT brings direct contingencies and indirect verbal processes to bear on the experiential establishment of greater psychological flexibility through acceptance, defusion, establishment of a transcendent sense of self, being present, values, and building expanding patterns of committed action linked to those values”* (Hayes et al. (2004)).

Therefore, ACT appears to be particularly timely and relevant in the context of this application. In ACT the goal is the development of *psychological acceptance*; that is, the willingness to let go of avoidance and contact one's immediate experience, even if that experience includes fear, anxiety, unpleasant thoughts, or trauma-related memories (Luoma et al. (2007)).

ACT also promotes the use of exercises or techniques in order to improve perspective-taking. These exercises help the patient to create an awareness of personal experiences without the content of those experiences defining the conceptualised self (Thompson et al. (2013)).

With regard to the use of those variables as components of the bot's context, the variables of trust, commitment, and acceptance have been given 5 possible values, ranging from *none* to *total*, and they are meant to increase or decrease depending on the interactions that the user and bot might have. In an attempt to simulate human behaviour, only some interactions will

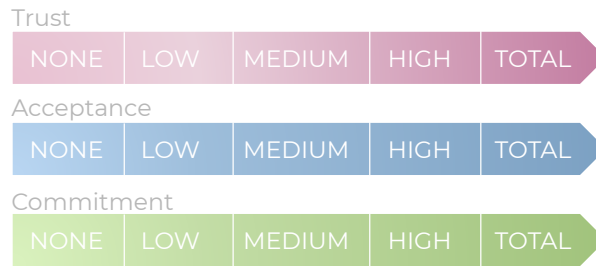


Figure 4.1: Ranges of possible values for the context variables.

Input	Trust	Acceptance	Commitment
“¿Que tal estás?”	Increase	-	-
“¿Lo pasas bien en clase?”	-	Increase	-
“¿Lo has contado a alguien?”	-	Increase	Increase
“¿Los otros niños se meten contigo?”	-	Increase	-
“Tu eres *”	Increase	-	-
“Te equivocas”	Increase	Increase	-
“¿Tienes amigos en clase?”	-	Increase	-
“No”	Increase	-	Increase
“Yo puedo ayudarte”	-	-	Increase
“Si”	Increase	-	Increase
“No te entiendo”	Increase	-	Increase

Table 4.1: User's input influence on bot's trust, acceptance, and commitment

increase or decrease the values of the variables and, for some interactions, they remain unaltered. Table 4.1 shows an excerpt of the inputs that affect the variables either by increasing or decreasing their values.

4.2.2. Implementation of the subject responses according to the current emotional state

As seen in Table 4.1, the user input can alter the bot's state, and that state will determine the kinds of responses given by the bot. That is, when the bot's trust is very low its answers will give less information about traumatic events that might have happened. Accordingly, if acceptance is high, the bot will be prone to taking advice from the user or accepting help from

him or her. Lastly, when commitment value is high or total, the bot will be open to asking for help and trying to do things in order to change its situation.

The following AIML code excerpt, shows an example of how the different values of commitment and acceptance combined produce different kinds of responses from the bot:

```
<category>
<pattern>^ TU CULPA</pattern>
<template>
  <condition name="acceptance" value="none">
    <condition name="commitment">
      <li value="none">a veces creo que la culpa la tengo yo</li>
      <li value="low">ellos dicen que es mi culpa, me gustaría ser diferente</li>
      <li value="medium">yo no hago nada malo a nadie</li>
      <li value="high">no es mi culpa, yo trato bien a los demás</li>
      <li value="total">sé que no es mi culpa, ¿qué puedo hacer?</li>
    </condition>
  </condition>
  <condition name="acceptance" value="low">
    <condition name="commitment">
      <li value="none">ellos dicen que es mi culpa, me gustaría ser diferente</li>
      <li value="low">ellos dicen que es mi culpa, me gustaría ser diferente</li>
      <li value="medium">sé que no es mi culpa, ¿qué puedo hacer?</li>
      <li value="high">sé que no es mi culpa, ¿qué puedo hacer?</li>
      <li value="total">no creo que sea mi culpa, a pesar de
        lo que digan los demás
      </li>
    </condition>
  </condition>
  <condition name="acceptance" value="medium">no creo que sea mi culpa,
    a pesar de lo que digan los demás</condition>
  <condition name="acceptance" value="high">
    <condition name="commitment">
      <li value="none">ellos dicen que es mi culpa, me gustaría ser diferente</li>
      <li value="low">no es mi culpa, yo no me meto con nadie</li>
      <li value="medium">sé que no es mi culpa, ¿qué puedo hacer?</li>
      <li value="high">¿crees que debería hablar con alguien del instituto?</li>
      <li value="total">no creo que sea mi culpa, a pesar de
        lo que digan los demás</li>
    </condition>
  </condition>
  <condition name="acceptance" value="total">sé que no es mi culpa,
    ¿qué puedo hacer?</condition>
</template>
</category>
```

In addition to the three context variables described, the use of AIML tags such as `<srai>` and `<that>` can also help conditioning OttoBot's responses depending on the context.

As described in Chapter 3 in Table 3.1 the `<that>` tag is an optional child of the category element that is used to establish the context of the pattern. If a category contains a `<that>` element, the pattern can only be matched if the last sentence of the bot's previous response matches the contents of that. Thus, the use of `<that>` makes the bot respond to a concrete request, if that request was in turn responding to a concrete bot response. For example:

```

<category>
  <pattern>Me llamo *</pattern>
  <that>Hola, ¿Cómo te llamas?</that>
  <template>
    <random>
      <li>Encantado de conocerte</li>
      <li>¡Qué bonito!</li>
      <li>Me gusta tu nombre</li>
      <li>Conozco a alguien con el mismo nombre</li>
    </random>
  </template>
</category>

```

In this example, the bot will only give one of those random responses when the user responds “*Me llamo **” (the asterisk is wildcard read with the `<star>` tag) to the message “*Hola, ¿Cómo te llamas?*”.

On the other hand, the `<srai>` tag allows the bot to recursively call categories, which allows the increment or decrement of the variables of trust, commitment, and acceptance in the response part of an interaction. In programming languages it is equivalent to calling a procedure from inside a function. The following example shows two categories, the first one is used when the bot receives “*¿Cómo te llamas?*” as input and, the second one is used inside the first one in order to increase OttoBot’s trust value after receiving such a message:

```

<category>
  <pattern>¿Cómo te llamas?</pattern>
  <template>
    <srai>INCREASE TRUST</srai>
    <random>
      <li>Me llamo Otto</li>
      <li>Otto, Y TÚ</li>
    </random>
  </template>
</category>

<category>
  <pattern>INCREASE TRUST</pattern>
  <template>
    <condition name="trust">
      <li value="none"><think><set name="trust">low</set></think></li>
      <li value="low"><think><set name="trust">medium</set></think></li>
      <li value="medium"><think><set name="trust">high</set></think></li>
      <li value="high"><think><set name="trust">total</set></think></li>
    </condition>
  </template>
</category>

```

4.3. Question set acquisition

In order to supply the bot with meaningful interactions for the topic of this study, research had to be conducted. The first bullying national study conducted in Spain was the AVE test (Piñuel and Oñate (2006)), which evaluated nearly 25.000 students across the country. This test was created to answer a need for school behaviour evaluation tools and it comprises the following diagnosis criteria:

- The existence of one or more aggressive behaviours.
- The repetition of such behaviour, which must also be acknowledged by the victim as something that he or she expects to happen inside the school context.
- The durability of the aforementioned behaviour, affecting the victim in several different aspects of his or her life (family, grades, friendship, etc.).

According to the AVE test, the coexistence of the above situations means that a child is being bullied.

Álvarez et al. (2013) have developed the CUVE³-EP and the CUVE³-ESO, which are two questionnaires that have been designed to identify how often students perceive different kinds of school violence committed by their peers. CUVE³-ESO was applied to 2597 compulsory secondary education students, from 18 high schools in Asturias and the results showed adequate psychometric properties. For this reason, this questionnaire has been chosen as the reference for the creation of the user-bot interactions rules regarding bullying diagnose.

Table 4.2 shows the relation of items composing the questionnaire that were relevant for the development of OttoBot. The complete original factors and items developed for the CUVE³-ESO can be found in Appendix B

4.3.1. Methodology

In order to increase OttoBot's chances of responding to the questions in real conversations, a survey was designed and distributed among students of the Complutense University of Madrid. The survey consisted in rewriting the 21 questions in a similar way, without altering its general meaning, but allowing subtle changes in meaning.

The survey was developed using Google Forms and the participants were contacted through the institutional email account of the university. The participants were given the instructions on how to complete the form both in the

Factor	Items
Peer-to-peer verbal violence	1. ¿Alguna vez se han dicho rumores sobre ti? 2. ¿Alguna vez han hablado mal de ti tus compañeros? 3. ¿Te han puesto algún mote molesto? 4. ¿Te han insultado tus compañeros?
Peer-to-peer physical violence and threats	5. ¿Has participado en alguna pelea en el instituto? 6. ¿Algún compañero te ha pegado? 7. ¿Alguna vez te han agredido cerca del instituto? 8. ¿Alguna vez te han amenazado para que hagas algo que no quieres? 9. ¿Alguna vez te han obligado a hacer algo que no quieres? 10. ¿Te han amenazado utilizando un arma? 11. ¿Algún compañero ha robado objetos del instituto? 12. ¿Te han robado algún objeto tus compañeros?
Social exclusion	13. ¿Te han excluido tus compañeros? 14. ¿Se han metido contigo por tus notas?
Violence through social networks and new technologies	15. ¿Te han amenazado públicamente a través de las redes sociales? 16. ¿Te han insultado públicamente en redes sociales? 17. ¿Han publicado alguna foto de ti sin tu permiso? 18. ¿Alguna vez han usado imágenes de ti para amenazarte? 19. ¿Te han enviado mensajes privados ofensivos? 20. ¿Te han amenazado a través de mensajes privados? 21. ¿Te han grabado alguna vez para burlarse de ti?

Table 4.2: CUVE³-ESO factors and items that compose it (Álvarez et al. (2013)).

email message and at the beginning of the form. The survey was distributed among 40 people from which only 15 did complete it. From the 15 responses, 4 had to be discarded because the participants actually answered the questions instead of rewriting them. Therefore, after the survey 11 versions of each of the questions were obtained. Both the message and the survey form itself can be found in the Apendix A.

Chapter 5

Formative evaluation

This chapter contains a detailed description of the phases conforming the evaluation process that has been carried out. Moreover, it also includes the relevant information regarding the profiles of the evaluators as well as the description and analysis of the results obtained from the mentioned evaluation.

According to Kendra and Rossett (2001), the evaluation is the process of examining a process or a program in order to determine what is working, what is not and why. Generally, evaluations can be divided into two sub-categories: formative and summative. A formative evaluation is a method for judging the worth of a program while the program activities are forming. This part of the evaluation focuses on the process and its goal is to gather feedback that can be used to improve the ongoing process. A summative assessment, on the contrary, is an evaluation that judges the worth of a program at the end of the program activities, that is, is focuses on the outcome. Due to the fact that OttoBot is an ongoing project that still needs to improve in its performance, a formative evaluation whose purpose is to examine the process was conducted. Given the limitations in the conversational agents state of the art, it is not possible to provide a complete system. A summative evaluation is, therefore, beyond the scope of this project.

OttoBot's responses are certainly limited to the number of requests or inputs. This makes it difficult for the bot to perform adequately in conversations. For this reason, one of the major limitations of OttoBot is the scarcity of input data examples that would make it much more responsive. The main goal of the formative evaluation is to establish to which extent the aforementioned lack of input limits OttoBot's performance as well as to acknowledge OttoBot's main improvement points. It must also be noted that this formative evaluation can provide as well more examples of the kind

of questions and requests that experts use in therapy, which would enrich OttoBot's input rules significantly.

5.1. Evaluation process

The evaluation process has been designed as a four-phase assessment method with the objective of observing to which degree the aforementioned issues limit OttoBot's performance.

- **First phase: unsupervised conversation.** In this phase, the professional has to chat with OttoBot on her own, during 20 minutes. The reason for choosing 20 minutes is that, after several informal tests with users it was observed that after 8 to 10 minutes either the users started losing interest or the conversations started to get stuck and going in circles. For this reason, and bearing in mind that the evaluation process is executed by professionals, it was decided to provide them with plenty of time for the conversation. Before the conversation starts, the psychologist is told to chat with Otto imagining it is a boy which suffers from school bullying and this is the first time they ever speak.
- **Second phase: assessment questionnaire.** After the unsupervised conversation, the professional is asked to complete a questionnaire regarding OttoBot's performance. The questionnaire, which can be seen in Figure 5.7, consists of 6 questions: 4 linear scale answer questions and 2 free-answer questions. The 4 linear-scale answer questions ask for the professional's appreciation on how OttoBot converses and the answer must say, on a scale from 1 to 10, being 1 "*not at all*" and 10 "*completely*", how much he or she agrees with the question. In addition, those questions also have a free writing space in which the professional must explain his or her answer. The last two questions ask the professional to give his or her opinion on the strong and weak points of OttoBot.
- **Third phase: supervised conversation.** In this phase, the professional has another 20-minute conversation with the bot but, in this case, the conversation must be supervised by the developer with the aim of giving him or her guidance in regard to word-choice and style. During this phase, the expert is explained what went wrong in the first conversation before starting the second one, and which are the technical reasons for that. Moreover, the context variables of trust, commitment and acceptance and the mechanisms that make them increase or decrease are explained. It must be noted that the aim of this phase is not to *dictate* the questions to the expert in order to make it

work, but to help him or her transform the sentence into a very similar one that can be understood by OttoBot.

- **Fourth phase: assessment questionnaire.** After the second conversation, the expert is given a text copy of the whole conversation with the aim of reading it as a whole. Then, the expert is asked to answer again the questionnaire from the second phase, which contains exactly the same questions, as depicted in Figure 5.7. In this phase, the expert is meant to give more detailed answers in the free-answer parts, bearing in mind the improvements that he or she has appreciated when comparing the first and second conversations.

5.1.1. Formative evaluator profiles

The reason for developing this project was to create a conversational agent that could emulate the behaviour of a 13 year-old boy that suffers school bullying with the objective of helping future psychologists in their training before real therapy. For this reason, the profiles needed for the assessment had to be qualified professionals with the corresponding studies in psychology.

The formative evaluation was carried out by two psychologists. The first subject (AF) has a Bachelor's Degree in Psychology and a Master's Degree in Psychopedagogy. In addition, she has extensive experience in working as a therapist with children and teenagers. The second subject (RA) has a Bachelor's Degree in Psychology and a Master's Degree in Art Therapy and Counseling.

5.2. Analysis of results

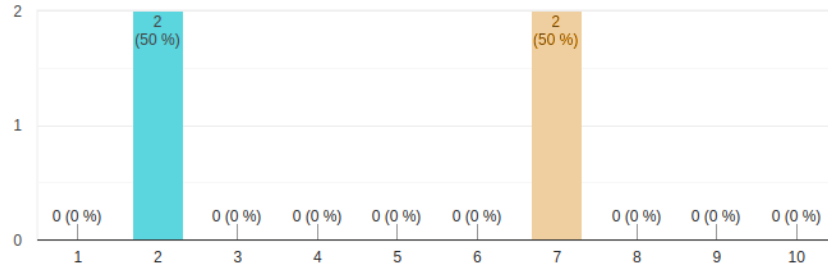
The first and third phases produced 2 conversations for each participant, which can be found in Appendices C and D. It is important to notice that when the bot is unable to find a response for a given request, it sends an ideogram (commonly known as *smileys* or *emojis*) of a thinking face, meaning that he does not know what to say to such a request.

Figure 5.1 shows the answers to the first question of the questionnaire: *"Is the conversation with OttoBot fluent?"*, in which both AF and RA gave a 2 over 10 in the second phase and a 7 over 10 in the fourth phase. In the image the blue bar and lines correspond to the answers given in the second phase while the brown bar and lines correspond to the answers given during the fourth phase.

AF answered *"He is not able to answer several questions in a row"* in the second phase and *"He is able to respond fluently"* in the fourth phase.

1. ¿Te ha parecido fluida la conversación con OttoBot?

4 respuestas



Por favor, explica tu respuesta:

4 respuestas

No consigue responder a varias preguntas seguidas
Consigue responder de manera bastante fluida
No entiende muchas cosas, por lo que es difícil que hable suficiente como para tener una conversación
Sabía contestar casi siempre y sus respuestas tenían sentido

Figure 5.1: Responses to question 1 given by the study subjects.

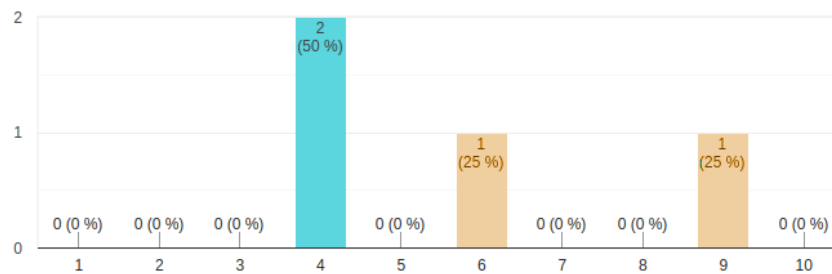
RA answered “*He does not understand many things, which makes it difficult to have a conversation with him*” in the second phase and “*He knew how to answer most of the times and the responses made sense*” in the fourth phase.

The results depicted in Figure 5.1 evidence indeed that, by lightening the fact that the input set is limited, the conversation is much more fluid and both experts changed their comments from a negative answer to a positive one, in which they highlight the fact that the bot is fluent and is able to react meaningfully to most of the conversation.

Figure 5.2 shows the answers to the second question of the questionnaire: “*Does OttoBot emulate a 13 year-old boy that suffers school bullying?*”, in which both AF and RA gave a 4 over 10 in the second phase and a 6 and a 9 over 10 in the fourth phase, respectively. In the image the blue bar and lines correspond to the answers given in the second phase while the brown bars and lines correspond to the answers given during the fourth phase.

2. ¿Consigue OttoBot simular a un adolescente que sufre acoso escolar?

4 respuestas



Por favor, explica tu respuesta:

4 respuestas

En ocasiones se atasca o no responde a la pregunta de forma adecuada
Los tiempos en comparación con un caso real en consulta son más rápidos.
En parte sí porque es como un adolescente que te cuenta un poco pero no quiere contarte todo y cuesta que te diga las cosas. Pero no entiende muy bien las conversaciones.
Parece que le cuesta hablar al principio pero enseguida te cuenta sus problemas y te pide ayuda

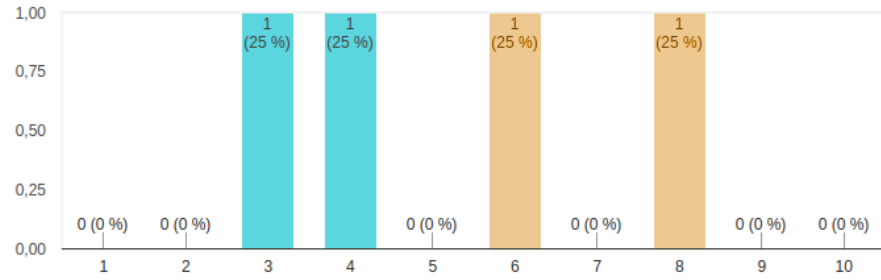
Figure 5.2: Responses to question 2 given by the study subjects.

AF answered “*Sometimes he gets stuck or does not respond adequately*” in the second phase and “*The timing compared to a therapy session in real life is faster*” in the fourth phase. RA answered “*In a way, he does because, like a teenager, he tells you a little bit but he does not want to tell everything, and it is hard for the therapist to get him to tell the things. But he does not understand the conversations well*” in the second phase and “*At the beginning it seems that it is hard for him to talk, but then he tells you his problems and asks for help*” in the fourth phase.

With regard to the results of question 2, in phase one, the expert AF highlights the fact OttoBot does not respond adequately, that is, it does not emulate the behaviour of a 13 year-old boy. However, in phase three, AM implies that it has improved, but the timing makes it less plausible because it usually takes several therapy sessions for the patient to tell the therapist

3. ¿Crees que las respuestas de OttoBot eran coherentes con el resto de la conversación?

4 respuestas



Por favor, explica tu respuesta:

4 respuestas

En ocasiones no da respuestas coherentes
Suele ser coherente salvo alguna pregunta
Cuando conseguía que entendiera lo que decía sí eran coherentes, pero casi nunca entendía lo que yo decía.
Sí, sus repuestas tienen mucho sentido y toda la conversación gira en torno al mismo tema

Figure 5.3: Responses to question 3 given by the study subjects.

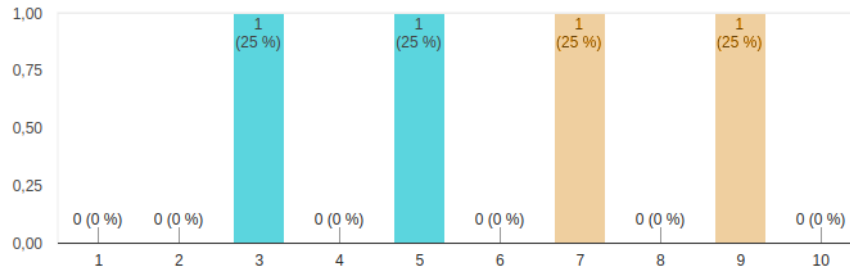
about difficult situations. RA, on the other hand, highlights the fact that OttoBot's limitations in understanding forces the therapist to think carefully about the questions and about which things will work for this concrete patient.

Figure 5.3 shows the answers to the third question of the questionnaire: “Do you think that OttoBot’s responses were coherent with the rest of the conversation?”, in which AF gave a 4 over 10 on the second phase and a 6 over 10 in the fourth phase. RA gave a 3 over 10 in the second phase and a 8 over 10 in the fourth phase. In the image the blue bars and lines correspond to the answers given in the second phase while the brown bars and lines correspond to the answers given during the fourth phase.

AF answered “*Sometimes the answers are not coherent*” in the second

4. ¿Te ha parecido que OttoBot puede servir de herramienta de aprendizaje a futuros profesionales?

4 respuestas



Por favor, explica tu respuesta:

4 respuestas

Podría ser una herramienta si se mejora
Puede ser de utilidad de cara a la formación práctica de profesionales
Puede funcionar para adaptarnos a alguien que no piensa como nosotros y cómo adaptar tu conversación a una persona que no te entiende, pero le falta fluidez
Tienes que adaptarte a su nivel lo que es muy parecido a un adolescente y tienes que pensar bien las preguntas que le haces

Figure 5.4: Responses to question 4 given by the study subjects.

phase and “*It is mainly coherent*” in the fourth phase. RA answered “*When he understood what I asked, the responses were coherent, but he almost never understood what I said*” in the second phase and “*Yes, the responses make a lot of sense and they revolve around the same topic*” in the fourth phase.

Coherence is one of the greatest limitations for chatbots. OttoBot also lacks coherence on conversation but the results on phase four show that, when supervising the conversation according to the methodology explained in section 5.1, there is more coherence in its responses. The experts RA and AF indicated this lack in their answers in phase 2, but both appreciated an improvement in the fourth phase of the experiment.

Figure 5.4 shows the answers to the fourth question of the questionnaire: “*Do you think that OttoBot can be of use as a tool for future professionals?*”, in which AF gave a 5 over 10 on the second phase and a 7 over 10 in the

5. ¿Qué aspectos de OttoBot destacarías como positivos?

4 respuestas

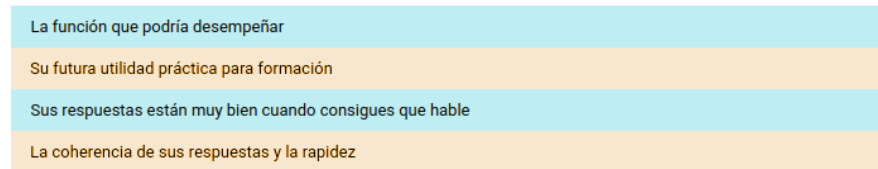


Figure 5.5: Responses to question 5 given by the study subjects.

fourth phase. RA gave a 3 over 10 in the second phase and a 9 over 10 in the fourth phase. In the image the blue bars and lines correspond to the answers given in the second phase while the brown bars and lines correspond to the answers given during the fourth phase.

AF answered *“It could be a tool with some improvements”* in the second phase and *“It can be useful for the practical training of professionals”* in the fourth phase. RA answered *“It can help us adapting to someone that does not think as we do, and how to adapt a conversation for someone that does not understand you, but it lacks conversational flow”* in the second phase and *“You have to adapt to his level, which makes him resemble a teenager and you have to think carefully the questions that you are going to make”* in the fourth phase.

The results obtained from question 4 indicate that, the experts believe that OttoBot could be a useful tool. In the second phase, AF conditions its usage as a training tool to an improvement while in the fourth phase, her answer is positive. RA, on the other hand, points out the characteristics that make OttoBot a good tool for the training of professionals in the future: it helps to adapt one’s thinking and conversation to someone that thinks differently and makes the therapist think deeply about what to say to the patient. Those two characteristics make therapy with OttoBot resemble a real therapy session with a 13 year-old boy suffering from school bullying.

Figure 5.5 shows the answers to the fifth question of the questionnaire: *“Which aspects of OttoBot would you highlight as positive?”*. In the image the blue lines correspond to the answers given in the second phase while the brown lines correspond to the answers given during the fourth phase.

AF answered *“The role it could play”* in the second phase and *“Its future practical use in education”* in the fourth phase. RA answered *“The answers are really well when you make him talk”* in the second phase and *“The co-*

6. ¿Qué aspectos crees que se deben mejorar?

4 respuestas

La fluidez en la conversación
Se podría avanzar más en la conversación y que pudiese contar e indagar más en el caso
El hecho de que no entiende el 90% de la conversación
El entendimiento de algunas de las partes de la conversación y la diversidad de temas a hablar con él

Figure 5.6: Responses to question 6 given by the study subjects.

herence of his answers and its speed” in the fourth phase.

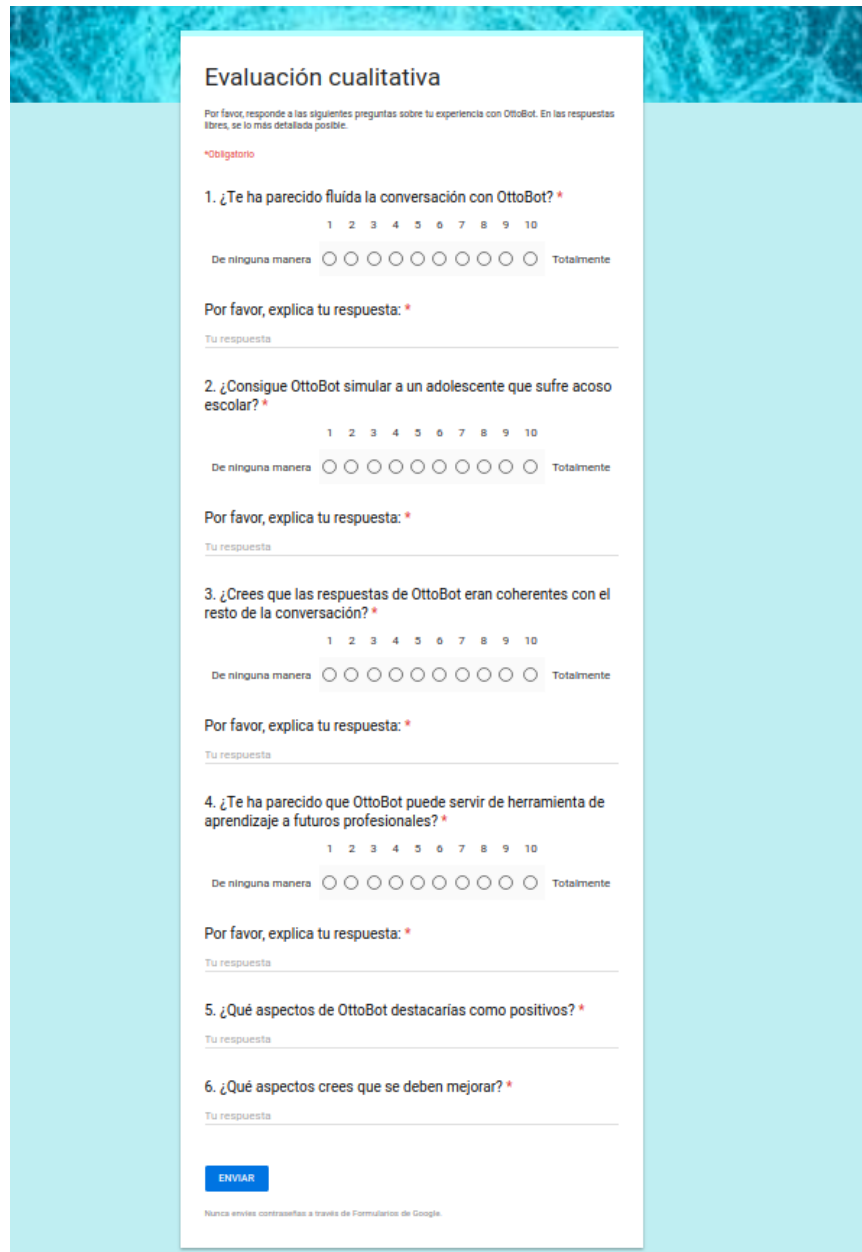
The results from question 5 are, in a sense, related to question 4 as well. The expert AF finds that the most positive aspect of OttoBot is the function it could have as a tool for training future psychologists, even in her answers from second phase. RA mentions coherence and speed as strong points of the conversational agent. These results indicate that, with the aforementioned improvements, OttoBot can be a very positive tool in order to help the future professionals of the field training for therapy.

Figure 5.6 shows the answers to the sixth question of the questionnaire: “Which aspects of OttoBot do you think need improvement?”. In the image the blue lines correspond to the answers given in the second phase while the brown lines correspond to the answers given during the fourth phase.

AF answered “*The conversational flow*” in the second phase and “*There could be some progress in conversation and he could tell more and get the therapist to know more about the case*” in the fourth phase. RA answered “*The fact that he does not understand 90% of the conversation*” in the second phase and “*Its understanding in some parts of the conversation and the diversity of topics when talking to him*” in the fourth phase.

The answers given to question 6 constitute the starting point of the improvement list developed for OttoBot. Fluidity in conversation, depth of the stories, greater understanding, and diversity of topics show the path to follow in order to enhance OttoBot’s capacities.

The overall view of the results described in this section is that, although OttoBot is not ready to be used in the present moment, it can be a useful tool for the training of future psychologists. The results indicate that the use of OttoBot could improve the users’ abilities in the treatment and therapy of children and teenagers suffering school bullying. The following chapter will discuss the results obtained from the formative evaluation.



Evaluación cualitativa

Por favor, responde a las siguientes preguntas sobre tu experiencia con OttoBot. En las respuestas libres, se lo más detallada posible.

***Obligatorio**

1. ¿Te ha parecido fluida la conversación con OttoBot? *

1 2 3 4 5 6 7 8 9 10

De ninguna manera ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Totalmente

Por favor, explica tu respuesta: *

Tu respuesta

2. ¿Consigue OttoBot simular a un adolescente que sufre acoso escolar? *

1 2 3 4 5 6 7 8 9 10

De ninguna manera ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Totalmente

Por favor, explica tu respuesta: *

Tu respuesta

3. ¿Crees que las respuestas de OttoBot eran coherentes con el resto de la conversación? *

1 2 3 4 5 6 7 8 9 10

De ninguna manera ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Totalmente

Por favor, explica tu respuesta: *

Tu respuesta

4. ¿Te ha parecido que OttoBot puede servir de herramienta de aprendizaje a futuros profesionales? *

1 2 3 4 5 6 7 8 9 10

De ninguna manera ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐ Totalmente

Por favor, explica tu respuesta: *

Tu respuesta

5. ¿Qué aspectos de OttoBot destacarías como positivos? *

Tu respuesta

6. ¿Qué aspectos crees que se deben mejorar? *

Tu respuesta

ENVIAR

Nunca envíes contraseñas a través de Formularios de Google

Figure 5.7: Assessment questionnaire.

Chapter 6

Discussion

In this chapter, the limitations extracted from the evaluation and analysis of the results in Chapter 5 will be acknowledged and discussed. Likewise, this chapter presents the reader with the identification and discussion of the positive features that have been achieved with the development of OttoBot.

The objective of this project was not to solve the conversational agent problem of understanding human context, but to provide the starting point for a prototype of a tool that will help future professionals of psychology in their training for therapy with children and teenagers suffering from school bullying. Therefore, this project has limitations that can establish the future work path in the improvement of such prototype.

One of the main limitations of this study is that OttoBot does not understand most of the requests from the user. This issue is due to the lack of input data in which the set of rules of OttoBot is based. In order to make OttoBot more responsive, an extensive quantity of input data would have to be retrieved and added to its core. As a matter of fact, this dataset could be retrieved from a greater survey in which many conversations would be processed. The problem of this approach is that therapy sessions are confidential and patients usually tell very intimate issues, which they do not wish to share, especially if they are under-age patients. Due to the nature of this kind of conversations, this would present itself as an extremely delicate matter both legally and ethically.

Another aspect that has been limited by OttoBot's performance is the evaluation. The characteristics of OttoBot have allowed the realization of a qualitative evaluation in order to establish not only the advantages and limitations of OttoBot but also the grounds for future improvements that can be achieved with further work on the subject. However, those characteristics did not allow for a quantitative evaluation that would provide the

system with greater quality assurance as well as a better understanding of its advantages and limitations.

The system created during the development of OttoBot could have been enriched by the use of a machine learning engine that could have helped OttoBot understanding context and inputs. Machine learning algorithms are optimised for raw, detailed source data. Thus, the data environment must provision large quantities of raw data for discovery-oriented analytics practices. However, as it has been mentioned earlier in this chapter, the lack of a large dataset of conversations in the topic precluded machine learning usage. In addition to the already discussed reasons, it must be noted that introducing machine learning in this project implies the processing of natural language, which is a limited field and a complicated task.

Another way of introducing machine learning in a chatbot system is to make it learn from the inputs that it receives. This technique is usually helpful when the conversational agent and the potential users share the same profile but, in this case, the users would be therapists while the bot emulates a teenager suffering school bullying. Thus, this approach cannot be applied to the improvement of OttoBot.

In spite of the limitations described throughout this chapter, relative merits have also been achieved in the development of OttoBot. First of all, the model that has been designed and developed can become a valuable base for future conversational models. The evaluation results endorse this assertion by selecting its possible future uses as one of the most remarkable features of OttoBot.

Another achievement that must be noticed is the creation of a system architecture which is sturdy and modular. The design and implementation of such an architecture enables the use of OttoBot with several different chat platforms as well as the addition of new technical resources as for example databases and machine learning tools without involving great architectural changes in the rest of the modules.

Finally, the methodology prototype proposed in this study focuses on an initial, simplified implementation, which is intended to progressively gain complexity and broaden its feature set. This methodology can therefore be used as the base for improving the actual system of OttoBot.

The questions discussed throughout this chapter provides confirmation that OttoBot is still a limited emotional model. However, the results obtained from the evaluations undeniably suggest that OttoBot serves a base for further meaningful explorations of the field.

Conclusions and Future Work

This project aimed to model a conversational agent that emulates the behaviour of a teenager that suffers from school bullying in order to train future psychologists. To achieve that objective, a prototype model has been designed and implemented in the form of a chat-based application.

Based on a qualitative evaluation of the system, it can be concluded that, although it is not ready at the moment, OttoBot can be of use as a tool for the training of future psychologists. The results indicate that the use of OttoBot can improve the users' abilities in the treatment and therapy of children and teenagers suffering school bullying.

The identification of the most relevant features of a patient's attitude that can influence the outcome of therapy has been carried out and a model for the state of the bot in relation to context has been created. However, the limitations of conversational agent technology have reduced the possibilities for the further development of the current project.

With regard to the technical achievements, a modular architecture in which each module can be easily replaced has been successfully developed and published as an online server. This architecture and deployment mode has provided OttoBot with both autonomy and maintainability and it has turned it into a flexible and open to change application.

7.1. Future work

The work presented in this project opens a number of potential avenues for future research. An increase of input data could be a great improvement for OttoBot's conversational capacities. Generating a large dataset would make it possible to add machine learning to the system architecture in order

to recognise the intent of the input from the users and then create responses based on a contextual analysis.

Another aspect that could be further developed is the granularity of the context variables of trust, acceptance and commitment. Adding new degrees to the context variables could result in more states as well as new combinations of the three variables, which would enable new possible states for the bot.

Additionally, the set of variables that constitute OttoBot's emotional state could be expanded by adding new dimensions. To perform this task, it would be necessary to work with a team of professional psychologists with experience in therapy of school bullying who can advise the creation and evolution of these variables.

With regard to the evaluation of OttoBot, as it has been acknowledged in Chapter 6, the lack of a quantitative assessment is a limitation of the system. Both quantitative and qualitative evaluations have weaknesses but they are compensated by the strengths of the other. Further qualitative assessment could be completed with the assistance of psychologist in order to improve the system in a process oriented way in which the experts would validate the bot's behaviour from subjective, but close perspective. Besides, quantitative evaluation carried out by future professionals of psychology would give the system verification in a reliable and objective way, which would make it population oriented and generalizable.

Conclusiones y trabajo futuro

Este proyecto tiene el objetivo de modelar un agente conversacional que emule el comportamiento de un adolescente que sufre acoso escolar para capacitar a futuros psicólogos para el tratamiento terapéutico. Para lograr este objetivo, se ha diseñado e implementado un prototipo de modelo en forma de aplicación basada en chat.

Basándose en la evaluación cualitativa del sistema, se puede concluir que, aunque todavía es pronto para ser utilizada con éxito, OttoBot puede ser útil como herramienta para la formación de futuros psicólogos. Los resultados indican que el uso de OttoBot podría mejorar las capacidades de los usuarios en el tratamiento y terapia de niños y adolescentes que sufren acoso escolar.

Se ha llevado a cabo la identificación de las características más relevantes de la actitud de un paciente que pueden influir en el resultado de la terapia y se ha creado un modelo de estado del bot en relación con el contexto. Sin embargo, las limitaciones de la tecnología de agentes conversacionales han reducido las posibilidades de un mayor desarrollo del proyecto actual.

Con respecto a los logros técnicos, se ha desarrollado una arquitectura modular en la que cada módulo se puede reemplazar fácilmente y se ha publicado con éxito como un servidor en línea. Esta arquitectura y modo de implementación han proporcionado autonomía y mantenibilidad a OttoBot, y lo han convertido en una aplicación flexible y abierta a posibles cambios en el futuro.

7.1. Trabajo Futuro

El trabajo presentado en este proyecto abre una serie de caminos potenciales para futuras investigaciones. El aumento de los datos de entrada podría

ser una gran mejora para las capacidades de conversación de OttoBot. La generación de un gran conjunto de datos permitiría agregar el aprendizaje automático a la arquitectura del sistema para reconocer la intención de los usuarios y luego crear respuestas basadas en un análisis contextual.

Otro aspecto que podría desarrollarse aún más es la granularidad de las variables de contexto: confianza, aceptación y compromiso. Agregar nuevos grados a las variables de contexto podría generar más estados, así como nuevas combinaciones de las tres variables, lo que permitiría nuevos estados posibles para el bot.

Además, el conjunto de variables que conforman el estado emocional de OttoBot podría ampliarse agregando nuevas dimensiones. Para la realización de esta tarea, es necesario trabajar con un equipo de psicólogos expertos en terapia con personas que han sufrido acoso escolar que puedan asesorar la creación y evolución de dichas variables.

Con respecto a la evaluación de OttoBot, como se ha reconocido en el Capítulo 6, la falta de una evaluación cuantitativa es una limitación del sistema. Tanto la evaluación cuantitativa como la cualitativa tienen debilidades, pero ambas se ven compensadas por las fortalezas de la otra. Se podría completar una evaluación cualitativa adicional con la ayuda de más psicólogos expertos para mejorar el sistema de manera orientada al proceso en la que los expertos validarían el comportamiento del bot desde una perspectiva subjetiva, pero cercana. Adicionalmente, una evaluación cuantitativa realizada por futuros profesionales de la psicología le daría al sistema una verificación fidedigna y objetiva, lo que haría que la aplicación estuviera más orientada a la población y fuera más generalizable.

Appendix A

Survey: *Reescritura de preguntas*

Por favor, intenta escribir las siguientes preguntas de otra manera. Puedes utilizar sinónimos aunque cambien sutilmente el significado de la frase (por ejemplo, cambiar “*compañeros*” por “*amigos*”). Si se te ocurren varias opciones puedes ponerlas separadas por comas.

Los cambios pueden ser sutiles o radicales. Ejemplos:

¿Te han puesto algún mote molesto? > *¿Alguna vez tus amigos te ponen apodosos ofensivos?*

¿Alguna vez se han dicho rumores sobre ti? > *¿Algún compañero ha dicho cosas falsas sobre ti?*

1. ¿Alguna vez se han dicho rumores sobre ti?
2. ¿Alguna vez han hablado mal de ti tus compañeros?
3. ¿Te han puesto algún mote molesto?
4. ¿Te han insultado tus compañeros?
5. ¿Has participado en alguna pelea en el instituto?
6. ¿Algún compañero te ha pegado?
7. ¿Alguna vez te han agredido cerca del instituto?
8. ¿Alguna vez te han amenazado para que hagas algo que no quieres?
9. ¿Alguna vez te han obligado a hacer algo que no quieres?

10. ¿Te han amenazado utilizando un arma?
11. ¿Algún compañero ha robado objetos del instituto?
12. ¿Te han robado algún objeto tus compañeros?
13. ¿Te han excluído tus compañeros?
14. ¿Se han metido contigo por tus notas?
15. ¿Te han amenazado publicamente a través de las redes sociales?
16. ¿Te han insultado publicamente en redes sociales?
17. ¿Han publicado alguna foto de ti sin tu permiso?
18. ¿Alguna vez han usado imágenes de ti para amenazarte?
19. ¿Te han enviado mensajes privados ofensivos?
20. ¿Te han amenazado a través de mensajes privados?
21. ¿Te han grabado alguna vez para burlarse de ti?

A.1. Distribution Message

¿Me ayudas a terminar mi TFG rellenando este formulario? El objetivo es reescribir 21 preguntas para decir lo mismo de otra manera.

No hay que responder a todas, cualquier contribución es buena. Las respuestas son totalmente anónimas, no hay respuestas incorrectas ni has de preocuparte por la ortografía y gramática y, si puedes, intenta que suene natural, aunque modifiques sutilmente el significado de la frase.

¡¡¡Muchas gracias de antemano!!!

Appendix **B**

Factores del CUVE3-ESO e ítems que los componen

Factores Ítems Violencia verbal del alumnado hacia el alumnado:

- Hay estudiantes que extienden rumores negativos acerca de compañeros y compañeras.
- Los estudiantes hablan mal unos de otros.
- El alumnado pone mote molesto a sus compañeros o compañeras.
- El alumnado insulta a sus compañeros o compañeras.

Violencia verbal del alumnado hacia profesorado:

- El alumnado habla con malos modales al profesorado.
- El alumnado falta al respeto a su profesorado en el aula.
- Los estudiantes insultan a profesores o profesoras.

Violencia física directa y amenazas entre estudiantes:

- El alumnado protagoniza peleas dentro del recinto escolar.
- Determinados estudiantes pegan a compañeros o compañeras dentro del recinto escolar.
- Algunos alumnos o alumnas protagonizan agresiones físicas en las cercanías del recinto escolar.

- Los estudiantes amenazan a otros de palabra para meterles miedo u obligarles a hacer cosas.
- Algunos alumnos amenazan a otros con navajas u otros objetos para intimidarles u obligarles a algo.

Violencia física indirecta por parte del alumnado:

- Ciertos estudiantes roban objetos o dinero del centro educativo.
- Ciertos estudiantes roban objetos o dinero de otros compañeros o compañeras.
- Algunos estudiantes roban cosas del profesorado.
- Algunos alumnos esconden pertenencias o material del profesorado para molestarle deliberadamente.
- Determinados estudiantes causan desperfectos intencionadamente en pertenencias del profesorado.

Exclusión social:

- Hay estudiantes que son discriminados por compañeros por diferencias culturales, étnicas o religiosas.
- Algunos estudiantes son discriminados por sus compañeros o compañeras por su nacionalidad.
- Determinados estudiantes son discriminados por sus compañeros o compañeras por sus bajas notas.
- Algunos estudiantes son discriminados por sus compañeros/as por sus buenos resultados académicos.

Violencia a través de las tecnologías de la información y de la comunicación:

- Algunos estudiantes ofenden, insultan o amenazan a otros a través de mensajes en Tuenti, Facebook,...
- Los estudiantes publican en internet fotos o vídeos ofensivos de profesores o profesoras.
- Hay estudiantes que publican en Tuenti, Facebook,... comentarios de ofensa, insulto o amenaza a otros.

- Los estudiantes publican en internet fotos o vídeos ofensivos de compañeros o compañeras.
 - Hay estudiantes que graban o hacen fotos a profesores o profesoras con el móvil, para burlarse.
 - Hay alumnos que graban o hacen fotos a compañeros/as con el móvil para amenazarles o chantajearles.
 - Ciertos estudiantes envían a compañeros/as mensajes con el móvil de ofensa, insulto o amenaza.
 - Hay estudiantes que envían mensajes de correo electrónico a otros con ofensas, insultos o amenazas.
 - Algunos estudiantes graban o hacen fotos a compañeros o compañeras con el móvil, para burlarse.
- Ciertos estudiantes publican en Tuenti, Facebook... ofensas, insultos o amenazas al profesorado. Disrupción en el aula:
- El alumnado dificulta las explicaciones del profesorado hablando durante la clase.
 - El alumnado dificulta las explicaciones del profesor/a con su comportamiento durante la clase.
 - Hay alumnado que ni trabaja ni deja trabajar al resto.

Violencia del profesorado hacia el alumnado:

- El profesorado tiene manía a algunos alumnos o alumnas.
- El profesorado tiene preferencias por ciertos alumnos o alumnas.
- El profesorado castiga injustamente.
- El profesorado ignora a ciertos alumnos o alumnas.
- El profesorado ridiculiza al alumnado.
- El profesorado no escucha a su alumnado.
- Hay profesores y profesoras que insultan al alumnado.
- El profesorado baja la nota a algún alumno o alumna como castigo.
- Ciertos profesores o profesoras intimidan o atemorizan a algún alumno o alumna.
- El profesorado amenaza a algún alumno o alumna.

Appendix C

Evaluation conversations: Subject 1

C.1. Phase 1: First conversation

ANDREA FUENTES : hola
OTTOBOT: Hola, ¿qué tal?
ANDREA FUENTES : bien y tu
OTTOBOT: supongo que bien
ANDREA FUENTES : ¿qué te gusta hacer?
OTTOBOT: 😊
ANDREA FUENTES : ¿tienes peleas con tus amigos?
OTTOBOT: Tengo a Ana y a algún compañero de clase
ANDREA FUENTES : ¿a qué te gusta jugar?
OTTOBOT: 😊
ANDREA FUENTES : ¿Qué haces con tus amigos?
OTTOBOT: Tengo a Ana y a algún compañero de clase
ANDREA FUENTES : ¿Te gusta ir al instituto?
OTTOBOT: no me gusta el instituto
ANDREA FUENTES : ¿Por qué?
OTTOBOT: 😊
ANDREA FUENTES : ¿por qué no te gusta el instituto?
OTTOBOT: Antes era divertido, pero ahora lo paso mal
ANDREA FUENTES : ¿Alguien te molesta?
OTTOBOT: 😊
ANDREA FUENTES : ¿Algún compañero te molesta?
OTTOBOT: 😊
ANDREA FUENTES : ¿Te hacen cosas que no te gustan?
OTTOBOT: 😊
ANDREA FUENTES : ¿tus compañeros son malos contigo?
OTTOBOT: 😊

ANDREA FUENTES : ¿tus compañeros se portan mal contigo?
OTTOBOT: 😞
ANDREA FUENTES : ¿Algún compañero te pega?
OTTOBOT: 😞
ANDREA FUENTES : ¿Tus amigos te insultan?
OTTOBOT: sí, cuando se meten conmigo lo paso mal
ANDREA FUENTES : ¿Y qué haces?
OTTOBOT: 😞
ANDREA FUENTES : ¿cuando se meten contigo qué haces?
OTTOBOT: 😞
ANDREA FUENTES : ¿has hablado con algún profesor?
OTTOBOT: 😞
ANDREA FUENTES : ¿se lo has contado a tus profesores?
OTTOBOT: Tengo miedo de contarlo y que se enteren
ANDREA FUENTES : ¿te llaman chivato?
OTTOBOT: 😞
ANDREA FUENTES : ¿que cosas te hacen?
OTTOBOT: 😞
ANDREA FUENTES : ¿qué cosas te hacen?
OTTOBOT: 😞
ANDREA FUENTES : ¿de qué tienes miedo?
OTTOBOT: 😞
ANDREA FUENTES : ¿Alguien te ayuda?
OTTOBOT: 😞
ANDREA FUENTES : ¿Algún amigo te ayuda?
OTTOBOT: 😞
ANDREA FUENTES : ¿con quién juegas en el recreo?
OTTOBOT: 😞
ANDREA FUENTES : ¿se lo has dicho a tus padres?
OTTOBOT: Muchas veces he estado a punto de hablar con mis padres y contarles todo
ANDREA FUENTES : ¿y por qué no se lo cuentas?
OTTOBOT: 😞
ANDREA FUENTES : ¿se lo has dicho a tus profesores?
OTTOBOT: Muchas veces he estado a punto de hablar con mis padres y contarles todo
ANDREA FUENTES : ¿te gustaría que te ayudasen?
OTTOBOT: 😞
ANDREA FUENTES : ¿necesitas ayuda?
OTTOBOT: 😞
ANDREA FUENTES : ¿te gustaría cambiar de instituto?
OTTOBOT: Antes me lo pasaba bien en clase, pero ahora mis compañeros me hacen pasarlo mal
ANDREA FUENTES : ¿qué haces cuando te hacen pasarlo mal?

OTTOBOT: 😊
ANDREA FUENTES : ¿Cuál es tu asignatura favorita?
OTTOBOT: 😊
ANDREA FUENTES : ¿Te gustaría hacer nuevos amigos?
OTTOBOT: Tengo a Ana y a algún compañero de clase
ANDREA FUENTES : ¿Cómo te puedo ayudar?
OTTOBOT: 😊
ANDREA FUENTES : ¿Desde cuando te pasa esto?
OTTOBOT: 😊
ANDREA FUENTES : ¿Hace mucho tiempo que te pegan?
OTTOBOT: la verdad es que sí, y creo que se pasan
ANDREA FUENTES : hay que hacer algo
OTTOBOT: 😊
ANDREA FUENTES : esto no puede seguir así
OTTOBOT: 😊
ANDREA FUENTES : ¿Por qué no se lo dices a tu profe?
OTTOBOT: 😊
ANDREA FUENTES : Seguro que te pueden ayudar
OTTOBOT: 😊

C.2. Phase 3: Second conversation

ANDREA FUENTES : Hola
OTTOBOT: Hola, ¿qué tal?
ANDREA FUENTES : ¿Cómo te llamas?
OTTOBOT: Otto, ¿y tú?
ANDREA FUENTES : Andrea
OTTOBOT: Me gusta tu nombre
ANDREA FUENTES : ¿Cuántos años tienes?
OTTOBOT: 13
ANDREA FUENTES : ¿Que cosas te gusta hacer?
OTTOBOT: 😊
ANDREA FUENTES : ¿Te gusta el deporte?
OTTOBOT: Depende de mi estado de ánimo
ANDREA FUENTES : ¿cómo te va el instituto?
OTTOBOT: Antes me gustaba ir a clase, pero ahora ya no
ANDREA FUENTES : ¿Se meten contigo?
OTTOBOT: sí, se meten conmigo y lo paso mal, me siento muy mal
ANDREA FUENTES : ¿se meten contigo tus compañeros?
OTTOBOT: sí, cuando se meten conmigo lo paso mal
ANDREA FUENTES : ¿Se lo has contado a alguien?
OTTOBOT: tengo miedo de contarle todo y que se enteren

ANDREA FUENTES : No tengas miedo

OTTOBOT: ¿Tú puedes ayudarme a arreglarlo?

ANDREA FUENTES : Si, puedo ayudarte si tú quieres

OTTOBOT: 😊

ANDREA FUENTES : Puedo ayudarte

OTTOBOT: No creo que puedas

ANDREA FUENTES : A otros chicos en tu misma situación he conseguido ayudarles

OTTOBOT: 😊

ANDREA FUENTES : ¿Tienes amigos en clase?

OTTOBOT: antes nos llevábamos todos bien, pero ahora se meten conmigo a diario

ANDREA FUENTES : ¿Cómo te hace sentir eso?

OTTOBOT: muy triste y enfadado

ANDREA FUENTES : ¿Se lo has contado a alguien?

OTTOBOT: Muchas veces he estado a punto de hablar con mis padres y contarles todo

ANDREA FUENTES : ¿Por qué no lo has hecho?

OTTOBOT: ¿Tú puedes ayudarme a arreglarlo?

ANDREA FUENTES : Yo puedo ayudarte

OTTOBOT: Si a ti te gusta

ANDREA FUENTES : Quiero ayudarte

OTTOBOT: ¿Realmente quieres eso?

ANDREA FUENTES : sí

OTTOBOT: Vale, si tú lo dices...

ANDREA FUENTES : Me gustaría ayudarte

OTTOBOT: ¿de verdad puedes ayudarme?

Appendix D

Evaluation conversations: Subject 2

D.1. Phase 1: First conversation

ROCÍO ARENILLAS: Hola
OTTOBOT: Hola, ¿qué tal?
ROCÍO ARENILLAS: muy bien y tu?
OTTOBOT: 😊
ROCÍO ARENILLAS: muy bien y tú qué tal?
OTTOBOT: 😊
ROCÍO ARENILLAS: estoy muy bien
OTTOBOT: 😊
ROCÍO ARENILLAS: cómo estas?
OTTOBOT: bien, ¿y tú?
ROCÍO ARENILLAS: bien
OTTOBOT: Guay
ROCÍO ARENILLAS: qué haces?
OTTOBOT: 😊
ROCÍO ARENILLAS: Qué has ehcho hoy?
OTTOBOT: 😊
ROCÍO ARENILLAS: Qué has hecho hoy?
OTTOBOT: 😊
ROCÍO ARENILLAS: ¿Quieres contarme algo?
OTTOBOT: 😊
ROCÍO ARENILLAS: ¿Cómo te llamas?
OTTOBOT: Otto, ¿y tú?
ROCÍO ARENILLAS: Chio
OTTOBOT: Encantado de conocerte
ROCÍO ARENILLAS: igualmente
OTTOBOT: 😊

ROCÍO ARENILLAS: ¿cuántos años tienes?
OTTOBOT: Tengo 13 años, ¿y tú?
ROCÍO ARENILLAS: 29
OTTOBOT: 😊
ROCÍO ARENILLAS: qué joven
OTTOBOT: 😊
ROCÍO ARENILLAS: te gusta el colegio?
OTTOBOT: Antes me gustaba ir a clase, pero ahora ya no
ROCÍO ARENILLAS: por qué no te gusta?
OTTOBOT: Mmmm no sé
ROCÍO ARENILLAS: ha pasado algo?
OTTOBOT: 😊
ROCÍO ARENILLAS: ha pasado algo en el colegio?
OTTOBOT: 😊
ROCÍO ARENILLAS: tienes amigos?
OTTOBOT: antes nos llevábamos todos bien, pero últimamente siempre se meten conmigo
ROCÍO ARENILLAS: ¿quién se mete contigo?
OTTOBOT: 😊
ROCÍO ARENILLAS: ¿tienes amigos?
OTTOBOT: antes nos llevábamos todos bien, pero últimamente siempre se meten conmigo
ROCÍO ARENILLAS: qué te dicen en el colegio?
OTTOBOT: 😊
ROCÍO ARENILLAS: te tratan mal en el colegio?
OTTOBOT: 😊
ROCÍO ARENILLAS: hablas con alguien?
OTTOBOT: 😊
ROCÍO ARENILLAS: ¿qué tal con tus padres?
OTTOBOT: 😊
ROCÍO ARENILLAS: ¿estás bien con tus padres?
OTTOBOT: 😊
ROCÍO ARENILLAS: tienes problemas en el colegio?
OTTOBOT: 😊
ROCÍO ARENILLAS: tienes problemas?
OTTOBOT: 😊
ROCÍO ARENILLAS: Te insultan en el colegio?
OTTOBOT: 😊
ROCÍO ARENILLAS: ¿que has hecho hoy?
OTTOBOT: 😊
ROCÍO ARENILLAS: has ido al colegio hoy?
OTTOBOT: 😊
ROCÍO ARENILLAS: tienes alguien con quien hablar?
OTTOBOT: A lo mejor hablar me ayuda

ROCÍO ARENILLAS: quieres decirme algo?
OTTOBOT: 😊
ROCÍO ARENILLAS: quieres hablar?
OTTOBOT: Creo que necesito tu ayuda
ROCÍO ARENILLAS: ¿cómo puedo ayudarte?
OTTOBOT: ¿de verdad puedes ayudarme?
ROCÍO ARENILLAS: ¡Claro!
OTTOBOT: 😊
ROCÍO ARENILLAS: puedo ayudarte
OTTOBOT: ¿En serio?
ROCÍO ARENILLAS: sí
OTTOBOT: Vale, si tú lo dices
ROCÍO ARENILLAS: ¿cómo te ayudo?
OTTOBOT: 😊

D.2. Phase 3: Second conversation

ROCÍO ARENILLAS: Hola
OTTOBOT: Hola, ¿cómo te llamas?
ROCÍO ARENILLAS: Chio
OTTOBOT: Me gusta tu nombre
ROCÍO ARENILLAS: muchas gracias
OTTOBOT: 😊
ROCÍO ARENILLAS: ¿cómo te llamas?
OTTOBOT: Otto, ¿y tú?
ROCÍO ARENILLAS: qué tal estás?
OTTOBOT: Me gusta tu nombre
ROCÍO ARENILLAS: que tal?
OTTOBOT: bien, ¿y tú?
ROCÍO ARENILLAS: bien
OTTOBOT: OK
ROCÍO ARENILLAS: qué te gusta?
OTTOBOT: 😊
ROCÍO ARENILLAS: cuál es tu comida favorita?
OTTOBOT: 😊
ROCÍO ARENILLAS: te gusta el deporte?
OTTOBOT: No estoy seguro
ROCÍO ARENILLAS: tienes amigos?
OTTOBOT: los de mi clase se meten conmigo a diario
ROCÍO ARENILLAS: ¿te gusta ir a clase?
OTTOBOT: Antes era divertido, pero ahora lo paso mal

ROCÍO ARENILLAS: por qué?
OTTOBOT: La verdad es que no sé
ROCÍO ARENILLAS: te han insultado tus compañeros?
OTTOBOT: Sí, algunas veces em llaman gafotas
ROCÍO ARENILLAS: ¿como te hace sentir eso?
OTTOBOT: muy triste y solo
ROCÍO ARENILLAS: se lo has contado a alguien?
OTTOBOT: ¿A quién debería contárselo?
ROCÍO ARENILLAS: a algún amigo
OTTOBOT: 😞
ROCÍO ARENILLAS: tienes amigos con quien hablar?
OTTOBOT: antes nos llevábamos todos bien, pero ahora se meten conmigo a diario
ROCÍO ARENILLAS: algún compañero te ha agredido?
OTTOBOT: Una vez, a la salida de clase, unos compañeros me empujaron, pero me fui corriendo
ROCÍO ARENILLAS: ¿se lo has contado a alguien
OTTOBOT: MUCHAS VECES HE ESTADO A PUNTO DE HABLAR CON MIS PADRES Y CONTÁRLES TODO
ROCÍO ARENILLAS: por qué no lo has hecho?
OTTOBOT: ¿Tú puedes ayudarme a arreglarlo?
ROCÍO ARENILLAS: puedo ayudarte
OTTOBOT: ¿De verdad puedes?
ROCÍO ARENILLAS: sí
OTTOBOT: Vale, si tú lo dices
ROCÍO ARENILLAS: qué tal con tu familia?
OTTOBOT: Mi familia no entiende lo que me pasa
ROCÍO ARENILLAS: ¿qué te pasa?
OTTOBOT: mis compañeros se meten mucho conmigo
ROCÍO ARENILLAS: juegas con tus compañeros?
OTTOBOT: antes nos llevábamos todos bien, pero ahora se meten conmigo a diario
ROCÍO ARENILLAS: sólo se meten contigo?
OTTOBOT: 😞

Bibliography

- ALHUMOUD, S., AL WAZRAH, A. and ALDAMEGH, W. Arabic chatbots: A survey. *International Journal of Advanced Computer Science and Applications*, Vol. 9(8), 535–541, 2018.
- ARORA, C. M. Defining bullying: Towards a clearer general understanding and more effective intervention strategies. *School Psychology International*, Vol. 17(4), 317–329, 1996.
- BARON, R. A. and RICHARDSON, D. R. *Human aggression*. Springer Science & Business Media, 2004.
- BOSWORTH, K., ESPELAGE, D., DUBAY, T., DAYTNER, G. and KARAGEORGE, K. Preliminary evaluation of a multimedia violence prevention program for adolescents. *American Journal of Health Behavior*, Vol. 24(4), 268–280, 2000.
- BOSZORMENYI-NAGY, I. and KRASNER, B. R. Trust-based therapy: a contextual approach. *The American Journal of Psychiatry*, 1980.
- BOULTON, M. J. Teachers’ views on bullying: Definitions, attitudes and ability to cope. *British Journal of Educational Psychology*, Vol. 67(2), 223–233, 1997.
- BRADEŠKO, L. and MLADENIĆ, D. A survey of chatbot systems through a loebner prize competition. In *Proceedings of Slovenian Language Technologies Society Eighth Conference of Language Technologies*, 34–37. 2012.
- CAMERON, G., CAMERON, D., MEGAW, G., BOND, R., MULVENNA, M., O’NEILL, S., ARMOUR, C. and MCTEAR, M. Towards a chatbot for digital counselling. In *Proceedings of the 31st British Computer Society Human Computer Interaction Conference*, 24. BCS Learning & Development Ltd., 2017.

- CANDELLO, H. and PINHANEZ, C. Designing conversational interfaces. *Simpósio Brasileiro sobre Fatores Humanos em Sistemas Computacionais-IHC*, 2016.
- CAREY, T. A. Improving the success of anti-bullying intervention programs: A tool for matching programs with purposes. *International journal of reality therapy*, Vol. 22(2), 2003.
- CARLSON, M., MARCUS-NEWHALL, A. and MILLER, N. Evidence for a general construct of aggression. *Personality and Social Psychology Bulletin*, Vol. 15(3), 377–389, 1989.
- CARPENTER, R. About the jabberwacky ai. 2011. [Online; accessed 30-June-2019].
- CHEN, H., LIU, X., YIN, D. and TANG, J. A survey on dialogue systems: Recent advances and new frontiers. *Acm Sigkdd Explorations Newsletter*, Vol. 19(2), 25–35, 2017.
- COLBY, K. M. Human-computer conversation in a cognitive therapy program. In *Machine conversations*, 9–19. Springer, 1999.
- CORCORAN, L., GUCKIN, C. and PRENTICE, G. Cyberbullying or cyber aggression?: A review of existing definitions of cyber-based peer-to-peer aggression. *Societies*, Vol. 5(2), 245–255, 2015.
- DESHPANDE, A., SHAHANE, A., GADRE, D., DESHPANDE, M. and JOSHI, P. M. A survey of various chatbot implementation techniques. *International Journal of Computer Engineering and Applications*, Vol. 11, 2017.
- ESBENSEN, F.-A. and CARSON, D. C. Consequences of being bullied: Results from a longitudinal assessment of bullying victimization in a multisite sample of american students. *Youth & Society*, Vol. 41(2), 209–233, 2009.
- FARRINGTON, D. P. Understanding and preventing bullying. *Crime and justice*, Vol. 17, 381–458, 1993.
- HALL, L., WOODS, S., DAUTENHAHN, K., SOBRAL, D., PAIVA, A., WOLKE, D. and NEWALL, L. Designing empathic agents: Adults versus kids. In *International Conference on Intelligent Tutoring Systems*, 604–613. Springer, 2004.
- HAYES, S. C., STROSAHL, K. D., BUNTING, K., TWOHIG, M. and WILSON, K. G. What is acceptance and commitment therapy? In *A practical guide to acceptance and commitment therapy*, 3–29. Springer, 2004.
- HENDRICKS, L., LUMADUE, R. and WALLER, L. R. The evolution of bullying to cyber bullying: An overview of the best methods for implementing

- a cyber bullying preventive program. In *National Forum Journal of Counseling and Addiction*, Vol. 1. 2012.
- HIGH, R. The era of cognitive systems: An inside look at ibm watson and how it works. *IBM Corporation, Redbooks*, 2012.
- HINDUJA, S. and PATCHIN, J. W. *Bullying beyond the schoolyard: Preventing and responding to cyberbullying*. Corwin Press, 2014.
- HUGHES, T. *Tom Brown's School Days: By an Old Boy*. Harper, 1870.
- HUSSAIN, S., SIANAKI, O. A. and ABABNEH, N. A survey on conversational agents/chatbots classification and design techniques. In *Workshops of the International Conference on Advanced Information Networking and Applications*, 946–956. Springer, 2019.
- KANETSUNA, T. and SMITH, P. K. Pupil insights into bullying, and coping with bullying: A bi-national study in japan and england. *Journal of School Violence*, Vol. 1(3), 5–29, 2002.
- KANTAR WORLDPANEL. Cuota de mercado de sistemas operativos en España. <https://es.kantar.com/tech/m%C3%B3vil/2019/enero-2019-cuota-de-mercado-de-smartphones-en-espa%C3%B1a/>, 2019. Accessed: 2019-08-25.
- KENDRA, S. and ROSSETT, A. Beyond the podium: Delivering training and performance to a digital world. 2001.
- KLEINKNECHT, R. A., DINNEL, D. L., KLEINKNECHT, E. E., HIRUMA, N. and HARADA, N. Cultural factors in social anxiety: A comparison of social phobia symptoms and taijin kyofusho. *Journal of anxiety disorders*, Vol. 11(2), 157–177, 1997.
- KOO, H. A time line of the evolution of school bullying in differing social contexts. *Asia Pacific Education Review*, Vol. 8(1), 107–116, 2007.
- LATHAM, A., CROCKETT, K. A., BANDAR, Z., FILIPE, J., FRED, A. and SHARP, B. A conversational expert system supporting bullying and harassment policies. 2010.
- LEE, M., ACKERMANS, S., VAN AS, N., CHANG, H., LUCAS, E. and IJSELSTEIJN, W. Caring for vincent: A chatbot for self-compassion. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 702. ACM, 2019.
- LUOMA, J. B., HAYES, S. C. and WALSER, R. D. *Learning ACT: An acceptance & commitment therapy skills-training manual for therapists*. New Harbinger Publications, 2007.

- MISHNA, F., SCARCELLO, I., PEPLER, D. and WIENER, J. Teachers' understanding of bullying. *Canadian Journal of Education/Revue canadienne de l'éducation*, 718–738, 2005.
- MONKS, C. P. and SMITH, P. K. Definitions of bullying: Age differences in understanding of the term, and the role of experience. *British Journal of Developmental Psychology*, Vol. 24(4), 801–821, 2006.
- MORENO, G. ¿hay alternativa a whatsapp? <https://es.statista.com/grafico/13743/hay-alternativa-a-whatsapp/>, 2018. Accessed: 2019-08-25.
- NANSEL, T. R., OVERPECK, M., PILLA, R. S., RUAN, W. J., SIMONS-MORTON, B. and SCHEIDT, P. Bullying behaviors among us youth: Prevalence and association with psychosocial adjustment. *Jama*, Vol. 285(16), 2094–2100, 2001.
- NOCENTINI, A., ZAMBUTO, V. and MENESINI, E. Anti-bullying programs and information and communication technologies (icts): A systematic review. *Aggression and Violent Behavior*, Vol. 23, 52–60, 2015.
- OED ONLINE. Bully, v.1. Oxford University Press. <https://www.oed.com/view/Entry/24610?> accessed 26 June 2019, 1989a.
- OED ONLINE. Bullying, n. Oxford University Press. <https://www.oed.com/view/Entry/24614> accessed 26 June 2019, 1989b.
- OED ONLINE. Violence, n. Oxford University Press. <https://www.oed.com/view/Entry/223638?rskey=c44mm4&result=1> accessed 27 June 2019, 1989c.
- OLWEUS. *Bullying at School: What We Know and What We Can Do*. Blackwell Publishers, 1993.
- OLWEUS, D., LIMBER, S. and MIHALIC, S. Blueprints for violence prevention, book nine: Bullying prevention program. *Boulder, CO: Center for the Study and Prevention of Violence*, 1999.
- PABIAN, S. and VANDEBOSCH, H. An investigation of short-term longitudinal associations between social anxiety and victimization and perpetration of traditional bullying and cyberbullying. *Journal of youth and adolescence*, Vol. 45(2), 328–339, 2016.
- PIÑUEL, I. and OÑATE, A. Test ave: Acoso y violencia escolar. *TEA EDICIONES*, 2006.
- RAMESH, K., RAVISHANKARAN, S., JOSHI, A. and CHANDRASEKARAN, K. A survey of design techniques for conversational agents. In *International*

- Conference on Information, Communication and Computing Technology*, 336–350. Springer, 2017.
- RIGBY, K. Consequences of bullying in schools. *The Canadian journal of psychiatry*, Vol. 48(9), 583–590, 2003.
- ROTTER, J. B. A new scale for the measurement of interpersonal trust 1. *Journal of personality*, Vol. 35(4), 651–665, 1967.
- SAFRAN, D. and SAFRAN, E. Creative approaches to minimize the traumatic impact of bullying behavior. *Creative interventions with traumatized children*, 132–167, 2008.
- SCHUSSLER, D., FRANK, J., LEE, T.-K. and MAHFOUZ, J. Using virtual role-play to enhance teacher candidates’ skills in responding to bullying. *Journal of Technology and Teacher Education*, Vol. 25(1), 91–120, 2017.
- TECKCHANDANI, N., SANTOKHEE, A. and BEKAROO, G. Aiml and sequence-to-sequence models to build artificial intelligence chatbots: insights from a comparative analysis. In *International Conference on Emerging Trends in Electrical, Electronic and Communications Engineering*, 323–333. Springer, 2018.
- TERRIBLENEWS. Aiml2.js. <https://github.com/TerribleNews/AIML2.js>, 2016.
- THOMPSON, B. L., LUOMA, J. B. and LEJEUNE, J. T. Using acceptance and commitment therapy to guide exposure-based interventions for post-traumatic stress disorder. *Journal of Contemporary Psychotherapy*, Vol. 43(3), 133–140, 2013.
- WEIZENBAUM, J. Eliza—a computer program for the study of natural language communication between man and machine. *Communications of the ACM*, Vol. 9(1), 36–45, 1966.
- YOUNG OH, E., SONG, D. and HONG, H. Interactive computing technology in anti-bullying education: The effects of conversation-bot’s role on k-12 students’ attitude change toward bullying problems. *Journal of Educational Computing Research*, 0735633119839177, 2019.
- ZHONG, H., LI, H., SQUICCIARINI, A. C., RAJTMAJER, S. M., GRIFFIN, C., MILLER, D. J. and CARAGEA, C. Content-driven detection of cyberbullying on the instagram social network. In *IJCAI*, 3952–3958. 2016.
- ZYCH, I., BALDRY, A. C. and FARRINGTON, D. P. School bullying and cyberbullying: Prevalence, characteristics, outcomes, and prevention. In *Handbook of behavioral criminology*, 113–138. Springer, 2017.

ÁLVAREZ, D., NÚÑEZ, J. and DOBARRO, A. Cuestionario para evaluar la violencia escolar en educación secundaria: CUVE3-ESO. 2013.

